

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Ches. Galy

Searcher Phone #: 308.4501 0

Searcher Location: Boston Lib.

Date Searcher Picked Up: 6/7/01

Date Completed: 6/7/01

Date Completed: 9/1/01

Searcher Prep & Review Time: _____

Clerical Prep Time: 1.2 hr

Type of Search

NA Sequence (#)_____

AA Sequence (#) 1

Structure (#) _____

Bibliographic

Litigation _____

Fulltext

Patent Family

Patent Family _____

(Other _____)

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr. Link _____

Lexis/Nexis

Sequence Systems ABSS02

WWW/Internet

Other (specify) _____

THIS PAGE BLANK (USPTO)

FT	/note= "WD motif"
FT	516..544
FT	/note= "WD motif"
PX	
PN	W09J98969-AI.
PX	
PD	05-AUG-1999.
PE	
PF	29-JAN-1999; 99WO-FR00196.
PX	
PR	09-DEC-1998; 98FR-0015545.
PR	30-JAN-1998; 98FR-0001100.
PX	
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE. (INSP) INST PASTEUR.
PI	Arenzana Seisdedos F, Benarous R, Concordet J, Durand H; Kroll M, Margottin F;
DR	WPI: 1999-469329/39. N-PDB: X86501.
PT	New human beta-transducin repeat containing protein and its fragments useful as, or to screen for, antiviral, antitumour, anti-inflammatory and anti-Alzheimer's agents
PS	Claim 1; Page 60-61; 71pp; French.
XX	
XX	The present sequence represents a human beta-transducin repeat containing protein (beta-Trcp). The protein directs proteins to the proteosome degradation pathways. The protein is able to interact with the Vpu protein of human immune deficiency virus-1 (HIV-1), cellular proteins Itapab or beta-catenin (bc) and/or protein Skp1. The protein controls ubiquitylation of phosphorylated proteins and thus their targeting to proteosomes for degradation. Depending on whether the process is inhibited or promoted, the result may be delayed breakdown of CD4 (in cases of HIV-1 infection); increased activity of IkB (and thus reduced activity of NFkappaB); increased degradation of mutant bc in tumour cells, or increased bc survival (and reduced apoptosis) in Alzheimer's patients. The beta-Trcp protein, and its active peptide fragments, or its nucleic acid, are used to screen for anti HIV-1 agents (antivirals), antitumour agents that disrupt cell cycle regulation or protein degradation in human tumour cells, and anti-inflammatory agents that disrupt activation of NFkappaB. Fragments of the protein are also useful for treating osteo-articular inflammation or acute inflammation associated with release of tumour necrosis factor.
XX	
Sequence	569 AA:
5Q	

	Query Match	Similarity	100.0%	Score	3034	DB	207	Length	569
	Best Local	Similarity	100.0%	Pred.	No. 2.4e-287				
	Matches	569	Conservative	0	Mismatches	0	Indels	0	Gaps
Qy	1	MDPAEALVLEKALKFMNSSEREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA	60						
Dd	1	mdpaevaiqelaklfmnssecedcngnepkllipeknslrqtnscarlclngetvcia	60						
Qy	61	STAKTEKCVAKTKLANGTSMIYPKRKLSASERKEKLCVKEFEQWSSDDVEFEHL	120						
Dd	61	stanktecvaktklangtssmiypkqrklasaeyekelcvvfyfeqwsdqefvehl	120						
Qy	121	ISQCHVQHGHINSYLYKPMLOLDEITLALPARGLDIIAENILSYLDASKSLCAELVCKEWT	180						
Dd	121	isgmchyqhghinslykpmliqrdfcltalparglidhaenilslayldaskslaaelvckew	180						
Qy	181	RVTSDGMLAKKRLIRRYMRTDLSLWGLAEKRGWGYLYKNNPPQGNAPNSFYALPYKII	240						
Dd	181	rvtsgdgmllkrlirrymrtdslwglaeerrgygyltknppdgngapnsfyalykpii	240						
Qy	241	ODIETISNNMGCHSLQRIHCHSETSKGYVCLQYDQKTVSGLRDNTIKIWDKNTLECK	300						
Dd	241	qdietisnnwgcghslqrlhcrseetskgyvclqyddqkivsglrdrntliklwdkntleck	300						

Oy	301	RILNGHGSVLCLOYDEVIITGSSDSVRVWQVNTGEMLNTLIIHC EAVLHJL RFNNGM	360
Oy	301	RILNGHGSVLCLOYDEVIITGSSDSVRVWQVNTGEMLNTLIIHC EAVLHJL RFNNGM	360
Db	301	rlilnghgsvlicqydeviitgssdsrvvwqvnbgemlntliihceavljhrfnnghm	360
Oy	361	WTCSKRSIAVWMA\$PTDITLRRLVGHRAAVNVDFD\$KXIV\$ASGDRFTIKWNT\$TC	420
Db	361	wtc\$ksr\$ia\$vwma\$ptd\$itlrrlrvlghraavnvdf\$dkxiv\$asgdrftikwnt\$tc	420
Oy	421	EFVRFLNLRHKKRGICLOYRDLRLVWSSSDNTIRLMDTECGACRLVEGHELVRCIRFDN	480
Db	421	efvrlflnrhkkrgiclqyrdlrvlwsssdntirlmdtecgacrlveghelevrcirfdn	480
Oy	481	KRIVSGAYDGKIRKWDLYAALDPAPAGTCLRTLVHSGRVRFLQDFEQIVSS\$HDT	540
Db	481	krivsgaydgkirkwdlyaaldpapagtcrlrtlvhsgvrflqdfefqivss\$hdtt	540
Oy	541	ILIMFELNDPAAOA\$EPSPSRRTTYISR	569
Db	541	ilimfelndpaaoa\$epspsrptyistr	569

RESULT	2	
B12813		
ID	B12813	standard; protein; 569 AA.
XX		
XX	AC	
XX	B12813;	
DT	27-NOV-2000	(first entry).
XX		
DE	Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.	
XX		
XX	Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;	
KW	beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;	
KW	gene therapy; colon cancer; beta-transducin repeat containing protein;	
KW	beta-TrCP.	
XX		
OS	Homo sapiens.	
PN	JF2000166542-A.	
XX		
PD	20-JUN-2000.	
XX		
PE	02-DEC-1998; 98JP-0343437.	
XX		
PR	02-DEC-1998; 98JP-0343437.	
XX		
PA	(KAGA-) KAGAKU GIJUTSU SHINKO JICHO DAN.	
XX		
DR	WPI: 2000-485550/43.	
DR	N-PSDB; A73132.	
XX		
PT	F-box protein of ubiquitin ligase SCF complex which promotes the	
XX	ubiquitination of Ikappab or beta-catenin	
PS	Claim 3; Page 10-12; 19pp; Japanese.	
XX		
CC	The present invention describes an F-box motif protein of ubiquitin	
CC	ligase SCF complex which promotes the ubiquitination of Ikappab or	
CC	beta-catenin and is constituted by Skp1 protein, Cull1 protein and a	
CC	complex (SCF complex) of F-box protein containing F-box motif and WD40	
CC	repeat motif and has the amino acid sequence of 45 residues (B12811)	
CC	or one of two 569 residue sequences (B12812, which is mouse ubiquitin	
CC	ligase FMD1 protein) and (B12813, which is human beta-transducin repeat	
CC	containing protein (beta-TrCP)). The F-box protein can be used for the	
CC	gene therapy of colon cancer by being recombined to a virus vector.	
XX		
XX	Sequence 569 AA;	
XX		
Query Match	100.0%;	Score 3034; DB 21; Length 569;
Best Local Similarity	100.0%;	Pred. NO. 2,4e-287;
Mismatches 569; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MDPAEAVLQEKALKFNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLNQEYVCLA 60
XX |||||||
PT 1 mdpaeanlqgkalkkfmsseredcnngpeprkiipeknslrqtnscaarclnqetvcia 60
PT |||||||
DB 61 STAMKTENCYAKTKRLANGTSSMIVPKORKLSASTYKEKELCVKFFQWSSDOVEYEHL 120
QY 61 stamktencvacktklangtssmivpkorklsasyekekelcvkyfeqwssdqyefvehl 120
XX |||||||
DB 61 stamktencvacktklangtssmivpkorklsasyekekelcvkyfeqwssdqyefvehl 120
XX |||||||
QY 121 ISOMCHYOHGHINSYLRKPMLORFITLAPRGIDHIAENTSLVDKSLCAAEVYCKEWY 180
XX |||||||
DB 121 isomchyqhghinsylkpmloqrdftalpargldhaenllsyldakslcaaelvckewy 180
XX |||||||
QY 181 RYVSDGMLMKKILIERWVRTDLSLRGLAERMGQYLFKNRPPGMAPNPSFYALYPKII 240
DB 181 rvsdgmmlwkkliernvrtclsivrglraerrwgqylfknkppgdnappnsfyralypkii 240
XX |||||||
QY 241 QDIETIESNMRCGRHSLSQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTECK 300
DB 241 qdietsenwrcgrhsqgrlhrchrsetskgyvclqyddqkivsglrndntlikwknlteck 300
XX |||||||
QY 301 RIITGHTGSVLCQYDERVLIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 riltghtgsvlclqydervliitgssdstvrwvntgemlntlihnceavlhrlfrnngmm 360
XX |||||||
QY 361 VTGSKDRSLAVWDMASPTDITLRRVLVGRRAVNVDFDDKYIVSASGDRTIKVMNTSTC 420
DB 361 vtcskdrslavwmasptdltlrrvlvghraavnvdfddkylvasagdrtlkvmntstc 420
XX |||||||
QY 421 EFPRTLNGHRRGIACIQYRDRLVYVSSSDNTIRLMDIEGACALRVLEGHELVRCIRFDN 480
DB 421 efvrtlnghrrgiaciqyrdrlvvsdsdntlrlmdiegcacalrvlegheelvrcirfdn 480
XX |||||||
QY 481 KRIVSGAYDCKIKVMDLVALDPRAPAGTCLRLTVLHSGRVRLORDEQIYSSSHDDT 540
DB 481 krivsgaydgkikvmdlvaaoldrapagtlcrltlvehsgrvrlrlqdefqivssshddt 540
XX |||||||
QY 541 ILIMDFLNDPAAQAEPPRSPRTTYTISR 569
DB 541 ilimdfindpaaqaeprrsprrtytysr 569
XX |||||||
RESULT 3
Y96697
ID Y96697 standard; Protein: 569 AA.
XX
AC Y96697;
XX
DT 26-SEP-2000 (first entry)
XX
DE Human beta-TTCP.
XX
KW E3 ubiquitin ligase; beta-TTCP; F-box; WD protein; I-kappa-B; inhibitor;
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
KW anti-inflammatory; immunosuppressive; cytosolic.
XX
OS Homo sapiens.
XX
PN WO200034447-A2.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US29371.
XX
PR 10-DEC-1998; 98US-0210060.
XX
PA (SIGN-) SIGNAL PHARM INC.
PA (YISS) YISSUM RES & DEV CO.
XX
PI Manning AM, Mercurio F, Amlt S, Ben-eriah Y, Davis M, Hatzudai A;
PI Lavon I, Yaron A;
XX
DR MPI; 2000-431294/37.
```

```
DR N-PSDB; A51229.
XX
PT Polypeptide enhancing phosphorylated IkappaB ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX
PS Claim 21: Page 72-74; 77pp; English.
XX
CC Human beta-TTCP, an F-box/WD protein family member, has been shown to
CC have homology to human E3 ubiquitin ligase (E3). E3 enhances
CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of
CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
CC degradation via the ubiquitin pathway is useful for identifying
CC modulators of this process for use in treating diseases associated with
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
CC the F-box results in a protein that functions as a dominant negative
CC molecule in vivo. Transient over-expression of delta-beta-TTCP (a
CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
CC activity. E3 and beta-TTCP can be used to modulate NF-kappa-B to treat
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
SQ
Sequence 569 AA:
Query Match 100.0%; Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.4e-287;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPAEAVLQEKALKFNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLNQEYVCLA 60
DB 1 mdpaeanlqgkalkkfmsseredcnngpeprkiipeknslrqtnscaarclnqetvcia 60
XX |||||||
QY 61 STAMKTENCYAKTKRLANGTSSMIVPKORKLSASTYKEKELCVKFFQWSSDOVEYEHL 120
DB 61 stamktencvacktklangtssmivpkorklsasyekekelcvkyfeqwssdqyefvehl 120
XX |||||||
QY 121 ISOMCHYOHGHINSYLRKPMLORFITLAPRGIDHIAENTSLVDKSLCAAEVYCKEWY 180
DB 121 isomchyqhghinsylkpmloqrdftalpargldhaenllsyldakslcaaelvckewy 180
XX |||||||
QY 181 RYVSDGMLMKKILIERWVRTDLSLRGLAERMGQYLFKNRPPGMAPNPSFYALYPKII 240
DB 181 rvsdgmmlwkkliernvrtclsivrglraerrwgqylfknkppgdnappnsfyralypkii 240
XX |||||||
QY 241 QDIETIESNMRCGRHSLSQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTECK 300
DB 241 qdietsenwrcgrhsqgrlhrchrsetskgyvclqyddqkivsglrndntlikwknlteck 300
XX |||||||
QY 301 RIITGHTGSVLCQYDERVLIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 riltghtgsvlclqydervliitgssdstvrwvntgemlntlihnceavlhrlfrnngmm 360
XX |||||||
QY 361 VTGSKDRSLAVWDMASPTDITLRRVLVGRRAVNVDFDDKYIVSASGDRTIKVMNTSTC 420
DB 361 vtcskdrslavwmasptdltlrrvlvghraavnvdfddkylvasagdrtlkvmntstc 420
XX |||||||
QY 421 EFPRTLNGHRRGIACIQYRDRLVYVSSSDNTIRLMDIEGACALRVLEGHELVRCIRFDN 480
DB 421 efvrtlnghrrgiaciqyrdrlvvsdsdntlrlmdiegcacalrvlegheelvrcirfdn 480
XX |||||||
QY 481 KRIVSGAYDCKIKVMDLVALDPRAPAGTCLRLTVLHSGRVRLORDEQIYSSSHDDT 540
DB 481 krivsgaydgkikvmdlvaaoldrapagtlcrltlvehsgrvrlrlqdefqivssshddt 540
XX |||||||
QY 541 ILIMDFLNDPAAQAEPPRSPRTTYTISR 569
DB 541 ilimdfindpaaqaeprrsprrtytysr 569
XX |||||||
RESULT 4
```

Y83041
 ID Y83041 standard; Protein: 569 AA.
 AC Y83041;
 XX
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein FBP-1.
 XX
 XX F-box protein: FBP; diagnosis: treatment; screening; agonist;
 KW antagonist; proliferative disorder; differentiative disorder;
 KW breast cancer; prostate cancer; ovarian cancer; cancer;
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
 KW inflammatory disorder; human.
 XX
 OS Homo sapiens.
 XX
 PN MO200012679-A1.
 XX
 XX
 PD 09-MAR-2000.
 XX
 PF 27-AUG-1999; 99MO-US19560.
 XX
 PR 28-AUG-1998; 98US-0098355.
 PR 03-FEB-1999; 99US-0118568.
 PR 15-MAR-1999; 99US-0124449.
 XX
 PA (UYNY) UNIV NEW YORK STATE.
 XX
 PI Chiaur DS, Pagano M, Latres E;
 XX
 DR WPI: 2000-256635/22.
 DR N-PSDB; Z93350.
 XX
 XX Novel nucleic acid for screening compounds useful for treating
 PT proliferative and differentiative disorders such as cancer and immune
 PT disorders comprises sequences encoding ubiquitin ligases
 XX
 PS Disclosure; Figure 3a; 245pp; English.
 XX
 CC Nucleic acids encoding substrate-targeting subunits of ubiquitin
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis
 CC of proliferative and differentiated related disorders by measuring
 CC FBP gene expression. Cells expressing such proteins or
 CC their fragments are useful for screening compounds. The compounds
 CC are agonists or antagonists, which are useful for treating a
 CC proliferative or differentiative disorder in a mammal such as
 CC breast, ovarian and prostate cancer and small cell lung carcinoma
 CC and also major opportunistic infections, immune disorders,
 CC cardiovascular diseases and inflammatory disorders. FBP protein,
 CC analogs, derivatives and their subsequences, anti-FBP antibodies
 CC are also useful in diagnosis of the disorders.
 XX
 SO Sequence 569 AA.

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 2.4e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEVLDOKALKFENSSREDCCNNGEPKRIIPKNSLRQYSCARLCLNQTVCIA 60
 DB 1 mdpaeavldekalkfennsseredccnngpeprkripknsrlrqytscaalcinqetvcia 60
 QY 61 STAMKTEKNCVARTKLANSSMTIVPKORLSAYEKEKLCVKEQWSESPQVEFVHL 120
 DB 61 stamktencvartklangsmtivpkqrkissayekekcvtkyeqwesqvefvehl 120
 QY 121 ISQMCHVGHGHTNSYLKPLQDFITALPARGLDHAENILSYLDAKSLCAELVCKEM 180
 DB 121 isqmchvghghinsylkplqdfitalparglidhaenilsyldakslcaaelvckey 180
 QY 181 RVTSGMLMKKLIERMVRTDSLWRGLAERRGGOYLFFKNKPPDGNA PPSFYRALYPKII 240

DB 181 rvtsgmlwkkliermvrtdslwrqlaerrwgqylfknkppdgnappsfyralypkii 240
 QY 241 ODITETSNMRCGRHSIORHCRSEPSKGYVCLQYDOKIVSGLRONTJIKMDKNTLECK 300
 DB 241 oditetsnmrcgrhsiorhcrsepskgyvclqyddqkivsglrbntjlkmdkntleck 300
 QY 301 RILGHTGSVLCQYDERVITGSSDSTVVRWMDVNTGEMLNTLIHHCFAVLHLRFNNGMA 360
 DB 301 rilghgtsvclqydervitgssdstvvrwmdvntgemlntlihnceavhlrfrnngm 360
 QY 361 VTGSKDSTIAVMDMASEPTDITLRRLVGHRAAVNVDFDDKYIVASGDRTIKWNFTSC 420
 DB 361 vtgskdstiavmdmasptdiltlravlghraavnvdfddkyivasgdrtilkwnftsc 420
 QY 421 EFVRTLNGHKGRTACLOYRRRLVSSSDNTIRLPMIEGACLRVLEGHELVRCIRPN 480
 DB 421 efvrtlnghkgrtacloyrrrlvsssdntirlpmiegacrlvleghelevrcirfn 480
 QY 481 KRIVSGAYDGKIKWMDLVAALDPRAPAGTLCRLTIVEHSGRVERLQDFEQIVSSSHDT 540
 DB 481 krivsgaydgkikwmdlvaaldprapagtlcrltlvehsgvrverlqdfefqvssshdt 540
 QY 541 ILIWDPLNDPAQAEPSPSRRTYTIISR 569
 DB 541 ilwdfplndpaqaepspsrtytysr 569

RESULT 5
 Y83250
 ID Y83250 standard; Protein: 569 AA.
 AC Y83250;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein hbeatricp.
 XX
 KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
 KW human.
 OS Homo sapiens.
 XX
 PN MO200022110-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99MO-US23705.
 PF 09-OCT-1998; 98US-0103787.
 PR
 PA (HARD) HARVARD COLLEGE.
 PI Zhou P, Howley P;
 DR WPI: 2000-317970/27.
 DR N-PSDB; Z93710.

Targeting degradation of polypeptide useful for treating cancer and
 PT other proliferative disorders, involves conjugating polypeptide with
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 PT compound
 XX
 PS Claim 9; Page 171; 185pp; English.
 XX
 CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 CC ligases) which can be used for the targeted degradation of a target
 CC polypeptide in vivo. Targeted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or

this is a false

XX Human cell signaling proteins useful for, e.g. diagnosing cell
 PT proliferative and inflammatory disorders
 XX
 PS Claim 1; Page 77-78; 90pp; English.
 XX
 CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded
 CC by cDNA obtained from Incyte clone 3339149 of COLAUC701 library. It is
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is
 CC found to be homologous to beta-transducin repeats containing
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as
 CC hybridisation probe for detecting CSIGP related sequences or allelic
 CC variants. Recombinant CSIGP can be produced in host cells by transforming
 CC them with genetically engineered vectors. Agonists or antagonists can be
 CC used in the treatment of cell proliferative and inflammatory disorders
 CC associated with decreased or increased CSIGP expression. CSIGP is used in
 CC the diagnosis, prevention or treatment of cell proliferative disorders
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.
 CC
 XX
 SQ Sequence 569 AA:

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 2.4e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPAEAVLQEKALFKFNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
 1 mdpaeavlgelakfkfnserecnngepprkilpeknslrtyngscarlclngelvtcla 60
 61 STAKTENCVAKTKLANGTSSMIVPKORKLASYEKELCVKFEQWSESDQVEVEHL 120
 61 stamktencvaktklangtssmivpkqrklssayekekecvkfeqwsedqvevehl 120
 121 ISQMCYHOGHINSYKLPMLQRFDTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 121 isqmcyhoghinsylkplmqrdftalparglidhaenilsyldakslcaaelvckewy 180
 121 isqmcyhoghinsylkplmqrdftalparglidhaenilsyldakslcaaelvckewy 180
 181 RVTSDGMLKKLIERMVRPDSLRGLAERRGQYLFKNKPPDGNAPPSFTRALYPKII 240
 181 rvtsgdmlkkliermvrpdsllrglaerrgqylfknkppdgnappsftralypkii 240
 181 rvtsgdmlkkliermvrpdsllrglaerrgqylfknkppdgnappsftralypkii 240
 241 QDIETIESNMRCGRHSIORHCRSETSKGYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300
 241 qdietiesnmrcgrhsiorhcrsetskgycloyddokivsglrdntikimdkntleck 300
 241 qdietiesnmrcgrhsiorhcrsetskgycloyddokivsglrdntikimdkntleck 300
 301 RILTGHTGSVLCLOYDERVIITGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 301 rilthtgsvclclgydervilitssdsyrvwdvntgemlntlilhceavhlhrlfngmm 360
 301 rilthtgsvclclgydervilitssdsyrvwdvntgemlntlilhceavhlhrlfngmm 360
 361 VTCSKRSIAVWDMASPTDITLRRLVGHRAAVNVDPDDKIVYASGRTIKWNTSTC 420
 361 vtcskrsiavwmdasptditlrrvlvghraavnvdpddkivyasgtrtikwntstc 420
 361 vtcskrsiavwmdasptditlrrvlvghraavnvdpddkivyasgtrtikwntstc 420
 421 EFVRLTNGHRGRIAGLQYDRVLVSSGSSDNTIRLMDIEGACLRVLEGHHEELVRCIRFDN 480
 421 efvrltnghrgrlgaclqydrvlvssgssdntirlmdiegcacrlrvleghheelvrcirfdn 480
 421 efvrltnghrgrlgaclqydrvlvssgssdntirlmdiegcacrlrvleghheelvrcirfdn 480
 481 KRIVSAGYGGKIKVMDLVLAADPRAPAGTLCRLTVEHSGRVRLQDFEQIVSSSHDDT 540
 481 krivsagygkikvmdlvlaadprapagtlcrltvehsgrvrlqdfefqvssshddt 540
 481 krivsagygkikvmdlvlaadprapagtlcrltvehsgrvrlqdfefqvssshddt 540
 541 IILWDFLNDPAQAEPSPSRITYISR 569
 541 iilwdflnmpaqaepspsrityistr 569
 541 iilwdflnmpaqaepspsrityistr 569

RESULT 7
 B12812
 ID B12812 standard; protein: 569 AA.
 XX
 AC B12812;

XX 27-NOV-2000 (first entry)
 DT
 XX
 DE Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.
 XX
 KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikapab;
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;
 KW beta-Trcp.
 XX
 OS Mus musculus.
 XX
 PN JP2000166542-A.
 XX
 PD 20-JUN-2000.
 XX
 PF 02-DEC-1998; 98JP-0343437.
 XX
 PR 02-DEC-1998; 98JP-0343437.
 XX
 PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 XX
 DR WPI: 2000-485550/43.
 DR N-PSDB; A73131.
 XX
 PT F-box protein of ubiquitin ligase SCF complex which promotes the
 PT ubiquitination of Ikapab or beta-catenin
 PS Claim 2; Page 9-10; 19pp; Japanese.
 XX
 CC The present invention describes an F-box motif protein of ubiquitin
 CC ligase SCF complex which promotes the ubiquitination of Ikapab or
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40
 CC repeat motif and has the amino acid sequence of 45 residues (B12811)
 CC or one of two 569 residue sequences (B12812, which is mouse ubiquitin
 CC ligase FWD1 protein) and (B12813, which is human beta-transducin repeat
 CC containing protein (beta-Trcp)). The F-box protein can be used for the
 CC gene therapy of colon cancer by being recombined to a virus vector.
 CC
 XX
 SQ Sequence 569 AA:

Query Match 98.8%; Score 2997; DB 21; Length 569;
 Best Local Similarity 98.6%; Pred. No. 9.8e-284;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

1 MDPAEAVLQEKALFKFNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60
 1 mdpaeavlgelakfkfnserecnngepprkilpeknslrtyngscarlclngelvtcla 60
 61 STAKTENCVAKTKLANGTSSMIVPKORKLASYEKELCVKFEQWSESDQVEVEHL 120
 61 stamktencvaktklangtssmivpkqrklssayekekecvkfeqwsedqvevehl 120
 61 stamktencvaktklangtssmivpkqrklssayekekecvkfeqwsedqvevehl 120
 121 ISQMCYHOGHINSYKLPMLQRFDTALPARGLDHAENILSYLDKSLCAAEVCKEWMY 180
 121 isqmcyhoghinsylkplmqrdftalparglidhaenilsyldakslcaaelvckewy 180
 121 isqmcyhoghinsylkplmqrdftalparglidhaenilsyldakslcaaelvckewy 180
 181 RVTSDGMLKKLIERMVRPDSLRGLAERRGQYLFKNKPPDGNAPPSFTRALYPKII 240
 181 rvtsgdmlkkliermvrpdsllrglaerrgqylfknkppdgnappsftralypkii 240
 181 rvtsgdmlkkliermvrpdsllrglaerrgqylfknkppdgnappsftralypkii 240
 241 QDIETIESNMRCGRHSIORHCRSETSKGYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300
 241 qdietiesnmrcgrhsiorhcrsetskgycloyddokivsglrdntikimdkntleck 300
 241 qdietiesnmrcgrhsiorhcrsetskgycloyddokivsglrdntikimdkntleck 300
 301 RILTGHTGSVLCLOYDERVIITGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 301 rilthtgsvclclgydervilitssdsyrvwdvntgemlntlilhceavhlhrlfngmm 360
 301 rilthtgsvclclgydervilitssdsyrvwdvntgemlntlilhceavhlhrlfngmm 360
 361 VTCSKRSIAVWDMASPTDITLRRLVGHRAAVNVDPDDKIVYASGRTIKWNTSTC 420
 361 vtcskrsiavwmdasptditlrrvlvghraavnvdpddkivyasgtrtikwntstc 420
 361 vtcskrsiavwmdasptditlrrvlvghraavnvdpddkivyasgtrtikwntstc 420

Db 361 vtcskdrsiawmdasptdlrlrrvlvghraavnvdfddkylvasagdrlikwntstc 420
 QY 421 EFWTLNGHKRGACIQYRDLVYSGSSDNTIRLMDIEGACLRVLEHETVRCIRPDN 480
 Db 421 efvrlngkrgiaciqyrdrlvsgssdntirlmdiegcacrlvleghetvrciridn 480
 QY 481 KRIYSGAVDGIKIKWDLVAALDPPAPAGTLCRLVHSGHVFRLQDFEQIVSSSHDDT 540
 Db 481 krlvsgaydgkikwldlmaaldpprapagtlcrlvlvshgsvfrlqdfefqivssshddt 540
 QY 541 ILIWDPLNDPAQAEPSPSRRTYTIYSR 569
 Db 541 lllwdfllndpaahaepprsrtytlysr 569
 RESULT 8
 Y83254
 ID Y83254 standard; Protein; 569 AA.
 AC Y83254;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein FWDLP.
 XX
 KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 KM destabilisation; proteolysis; drug discovery; gene therapy; cancer;
 KM oncoprotein; Huntington's disease; gene knockout; delivery systems;
 KM mouse; ss.
 XX
 OS Mus musculus.
 XX
 PN NC0200022110-A2.
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US23705.
 XX
 PR 09-OCT-1998; 98US-0103787.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Zhou P, Howley P;
 XX
 DR WPI; 2000-317970/27.
 DR N-PSDB; 293714.
 XX
 PT Targeting degradation of polypeptide useful for treating cancer and
 PT other proliferative disorders, involves conjugating polypeptide with
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 PT compound
 XX
 PS Claim 9; Page 184-185; 185pp; English.
 XX
 CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 CC ligases) which can be used for the targeted degradation of a target
 CC polypeptide in vivo. Targeted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or
 CC increasing the level of a target polypeptide and for creating and
 CC expressing a destabilized polypeptide which is subjected to SCF
 CC mediated proteolysis. Degrading any desired protein in a cell is
 CC useful for preventing or treating diseases caused by the presence of
 CC abnormal amount of the specific polypeptides, for drug discovery and
 CC for gene therapy. Diseases treated include cancer, by degradation of
 CC oncoproteins; Huntington's disease, other proliferative disorders and
 CC microbial infections. The method provides a quick and easy
 CC alternative to gene knockout technology. The target polypeptide can
 CC be degraded at all stages, or a specific stage, of development in the
 CC mature animal.
 XX
 SQ Sequence 569 AA;

Query Match 98.8%; Score 2997; DB 21; Length 569;
 Best Local Similarity 98.6%; Pred. No. 9 8e-284;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MDPAAVLQERALKFMNSEREDCNNEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
 Db 1 mdpaaavlqeralkfmnserecdnnepprklipeknsrlqtynscaclnoetvcclt 60
 QY 61 STAKTENCVAKTLANGTSSMIYPKQRKLSASTEKELCVKTFEQMSDESDYEFVEHL 120
 Db 61 staktencvaktlangtssmiypkqrklstasayekelcvkfefqmsesdyefvehl 120
 QY 121 ISOMCHVOHGHINSYLKPMLQDFETLAPARGLHINENISYLDARSLCAELVCKEMY 180
 Db 121 isgmchvghghinsylkpmldfetalparglhienlisyldarslcaaelvckemy 180
 QY 181 RVTSQGLMKKLIBRMVWRTDSLWRGLAERRGCGYLFRKNPPDGNAPPNSFYRALYPRKI 240
 Db 181 rvtsqglmkklbrmvwrtdslwrglaerrgcyglfrknppdgnappnsfyralypkii 240
 QY 241 QDFTIESNMRGSRHSIQRIHCRSETSKGYVCLQYDQKIVSGRLDNTIKIMDKNTLECK 300
 Db 241 qdftiesnwrctgrhsilqrhcrsetskgyvclqyddqkivsgrlrntikimdkntleek 300
 QY 301 RILGHGVSVCLOQYDERVITITGSSDSTVRVADYNTGEMLNTLIHGEAVYLHFRNNGMM 360
 Db 301 rilghgvsyvcloqydevvittgssdstvrvadynvtnagemlntlihgeavylhfrnngmm 360
 QY 361 VTCSKDRSIAVWMDASPTDITLRRVLVGHRAAVNVDFDDKYIVSASGDRTIKWNTSTC 420
 Db 361 vtcskdrsiawmdasptdlrlrrvlvghraavnvdfddkylvasagdrlikwntstc 420
 QY 421 EFWTLNGHKRGACIQYRDLVYSGSSDNTIRLMDIEGACLRVLEHETVRCIRPDN 480
 Db 421 efvrlngkrgiaciqyrdrlvsgssdntirlmdiegcacrlvleghetvrciridn 480
 QY 481 KRIYSGAVDGIKIKWDLVAALDPPAPAGTLCRLVHSGHVFRLQDFEQIVSSSHDDT 540
 Db 481 krlvsgaydgkikwldlmaaldpprapagtlcrlvlvshgsvfrlqdfefqivssshddt 540
 QY 541 ILIWDPLNDPAQAEPSPSRRTYTIYSR 569
 Db 541 lllwdfllndpaahaepprsrtytlysr 569
 RESULT 9
 R85852
 ID R85852 standard; Peptide; 517 AA.
 AC R85852;
 XX
 DT 13-SEP-1996 (first entry)
 XX
 DE WD-40 domain-contg. beta-TRCP protein.
 XX
 KM WD40 repeat region; beta-transducin; protein-protein interaction; drug;
 KM intracellular signalling; protein kinase C; homology; motif; modulator;
 KM receptors of activated protein kinase; enzyme activity; Isozyme; human.
 XX
 OS Synthetic.
 OS
 PN NC09521252-A2.
 PD 10-AUG-1995.
 XX
 PF 31-JAN-1995; 95WO-US01210.
 XX
 PR 01-FEB-1994; 94US-0190802.
 XX
 PA (STRD) UNIV LEELAND STANFORD JUNIOR.
 XX

PI Mochly-Rosen D, Ron D;
XX
DR WPI: 1995-283772/37.
XX
PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the
PT activity of a protein, eg. protein kinase C, which interacts with a
PT protein contg. a WD-40 region.
XX
PS Example 5; Page 80-82; 351pp; English.
XX
CC Proteins R85851-92 are protein which contain at least one WD-40 (also
CC called beta-transducin homologous) amino acid repeat motifs. The WD-40
CC regions are involved in protein-protein interactions between proteins
CC involved in intracellular signaling. An example of such an interaction
CC is between protein kinase C and receptors of activated protein kinase
CC (RACK), esp. RACK-1 (R85850). Proteins R85851-82 were isolated based on
CC homology with beta-transducin, whereas proteins R85882-92 were isolated
CC based on homology with the WD-40 consensus sequence (R85893). The
CC proteins were used to construct the peptides R84928-R85063 and
CC R85786-R85842. The peptides can be used to identify target proteins
CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of
CC proteins involved in protein-protein interaction and to screen for drugs
CC that will affect protein-protein interaction involving WD-40 domains.
XX
SQ Sequence 517 AA;

Query Match 85.1%; Score 2582.5; DB 16; Length 517;
Best Local Similarity 91.4%; Pred. No. 2,3e-243;
Matches 467; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

QY 18 SSREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTEVCATKLAN 77
DB 13 aserecdnngpprkiipteknlrtq-----tklan 42
QY 78 GISSMIVPKRKISASYEKEKELCVKFEQWESDQVEFEHLISOMCHYGHINSYLK 137
DB 43 gtsmivpqrkisanyekelcvkyfegwescdqvfevhlismchyghnitylk 102
QY 138 PMQORFIALPARGDHAENILSYLDAKSLCAELVCKEYRPTSDMLKKILERNV 197
DB 103 pmqorftalpargidhaenilslldakslcsaelvckeyrvtsdgmllkklernv 162
QY 198 RTDSLWGLAERKGMQVYFKNPPDGNAPNSFYRALYPKIIDIETIESNMGRHSLS 257
DB 163 rtdslwrglaerkgvqyfknpkpdgktrpnsfyralkyldietiesnmwrgihsl 222
QY 258 QRHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTSVLCIQYDE 317
DB 223 qrhcrsetskgyyclqyddqkivsglrdntlikwckntleckrylmghtsvlclqyde 282
QY 318 RVTITGSSDSTVWVMDVNTGEMNLIIHCEAVLHREFNNGMMVMYCSKDRSTAVVDMASP 377
DB 283 rvtitg-sdstvrvwdvntgemnlilnceavlhrlfngmmvtcskdrstlavvdmasa 341
QY 378 TDTLRLRVLVGHRAAVNVVDFDKYIVSASGDRITIKVNTSTCEPFRILNGHKGRIACIQ 437
DB 342 tdtlrlrvlvghraavnvdfdkyivsasgdrilkwntstcefrtlngkhkgriaciq 401
QY 438 YRRLVYSSSDNTIRLMDIEGACLRVLEGHEBLVRCIRFNNKRIIVSAGYGRKRWDL 497
DB 402 yrrlvysgsdntlrlmdiegacrlvlegheblvrcirfnnkriivsgaygkrlkvwdl 461
QY 498 VAALDRAPAGTCLRTLVHSGRVRLOFDEFQIVSSSHDPTIILMDLNP 550
DB 462 vaaldrapagtcrltlvhsgrvrlqfdeqivssshdptililwdfldnp 514

RESULT 10
ID Y96696
XX Y96696 standard; Protein: 542 AA.
AC Y96696;

XX 26-SEP-2000 (first entry)
DT Human E3 ubiquitin ligase.
XX
DE Human E3 ubiquitin ligase.
KW E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
KW anti-inflammatory; immunosuppressive; cytostatic.
XX
OS Homo sapiens.
XX
PN W0200034447-A2.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US29371.
XX
PR 10-DEC-1998; 98US-0210060.
XX
PA (SIGN-) SIGNAL PHARM INC.
PA (YISS) YISSUM RES & DEV CO.
PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
PI Lavon I, Yaron A;
DR WPI: 2000-431294/37.
DR N-PSDB; A51228.
XX
PT Peptide enhancing phosphorylated I-kappaB ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX
PS Claim 1; Page 70-72; 77pp; English.
XX
CC This is human E3 ubiquitin ligase (E3), which is homologous to human
CC beta-TrCP, an F-box/WD protein family member. E3 enhances ubiquitination
CC of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor
CC kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the
CC ubiquitin pathway is useful for identifying modulators of this process
CC for use in treating diseases associated with activation of NF-kappa-B. In
CC vitro analysis suggests that deletion of the F-box results in a protein
CC that functions as a dominant negative molecule in vivo. Transient
CC over-expression of delta-beta-TrCP (a deletion mutant) inhibited the
CC degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells,
CC resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be
CC used to screen for modulators of NF-kappa-B activity. E3 and beta-TrCP
CC can be used to modulate NF-kappa-B to treat inflammatory diseases,
CC autoimmune diseases, cancer and viral infections.
XX
SQ Sequence 542 AA;

Query Match 78.6%; Score 2384.5; DB 21; Length 542;
Best Local Similarity 79.0%; Pred. No. 5,8e-224;
Matches 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;

QY 1 MDPKAVLQEKALKFENSSERECNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 60
DB 1 mep-dsylvldtlelms-----vp-----tslwgcanlv---esmcsl 36
QY 61 S-TAMKTEQVAKTKLANGTSSMIVPKRKLSASYEKEKELCVKFEQWESDQVEFEV 118
DB 37 sclgmspsvcl---qslngtsvslvrkprpsengykdldklyfdqysedqeyfe 93
QY 119 HLISOMCHYGHINSYLKFMLODFETLALPARGDHAENILSYLDAKSLCAELVCKE 178
DB 94 hlismchyghninsylkfmldfetalpargdhaenilslldakslcsaelvcke 153
QY 179 MYRVTSDGMLAKKILERNVRTDSLWRGLAERGMGOVLFKNKPPDGNAPNSFYRALYRK 238
DB 154 wgrvtsgmlwklklierwrtdslwrglaergmgoylvfknkppdgnapnsfyralky 211

QY 239 IIDDITFESNMRCGRHSIQRIHCRSETSKGVYCLQYDQKIVSGLDNDTIKWKNTLE 298
Db 212 IIDDITFESNMRCGRHSIQRIHCRSETSKGVYCLQYDQKIVSGLDNDTIKWKNTLE 271
QY 299 CKRLTHTGTSVLCLOYDERVITITSSDSTVRVWDMVTGEMLNTLIHCEAVLHRLRNNG 358
Db 272 CLKVLGHSVSLQYDERVITITSSDSTVRVWDMVTGEMLNTLIHCEAVLHRLRNNG 331
QY 359 MMYTSCDRSIAVWDMASPTDILRLRVLGHRRAVNVDDDKIVSASGDRITKWNVS 418
Db 332 LMTCSKDSIAVWDMASPTDILRLRVLGHRRAVNVDDDKIVSASGDRITKWNVS 391
QY 419 TCEPVRFLNGHKRGICLQYDRDLVWSSGSDNTIRLMDIECGACRLVLEGHLEIVRCIRF 478
Db 392 TCEPVRFLNGHKRGICLQYDRDLVWSSGSDNTIRLMDIECGACRLVLEGHLEIVRCIRF 451
QY 479 DNRIVSAGVADGRIKWDVLAALDPRAPAGTLCRLTVEHSGRVERLOPFOIVSSSD 538
Db 452 DNRIVSAGVADGRIKWDVLAALDPRAPAGTLCRLTVEHSGRVERLOPFOIVSSSD 511
QY 539 DTLIMDLNDPAQAQAEPPSPSKRTTYTISR 569
Db 512 DTLIMDLNDPAQAQAEPPSPSKRTTYTISR 542

RESULT 11

Y22469 Y22469 standard; Protein; 626 AA.
AC Y22469;
DT 29-SEP-1999 (first entry)
DE Human gmyc-N-sel-10 protein sequence.
XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KM therapy; mammary gland.
XX Homo sapiens.
OS Synthetic.
XX MO9932623-A1.
XX 01-JUL-1999.
XX 17-DEC-1998; 98WO-US26820.
XX 19-DEC-1997; 97US-0068243.
XX (PHAA) PHARMACIA & UPJOHN CO.
XX Gurney ME, Li J, Pauley AM;
XX WPI: 1999-458026/38.
XX N-PSDB; X99712.
XX New isolated human sel-10 polypeptides
XX Example 3; Page 74-77; 91pp; English.
XX This sequence is an epitope-tagged version of a human sel-10 protein
XX of the invention. The human sel-10 proteins of the invention are isolated
XX from hippocampus and mammary gland. The polypeptides can be used to
XX alter presenilin function. Compounds which inhibit either the expression
XX or the activity of the human sel-10 polypeptides may reverse the effects
XX of mutations to presenilin-1 (PS-1) or PS-2, and therefore may be useful
XX for the prevention or treatment of Alzheimer's disease.
XX Sequence 626 AA;

Query Match 21.2%; Score 644.5; DB 20; Length 626;
Best Local Similarity 30.1%; Pred. No. 6e-54;

Matches 169; Conservative 103; Mismatches 222; Indels 67; Gaps 15;
QY 8 LOEKALFNMNSEREDCNENPPRIIPK--NSLRO--TVNSCARLCLNOETVCLASTA 63
Db 53 meqkllseednemesygdltmeqkllseedlnsmkklhdgsevrfsyqkpkcksey 112
QY 64 MKTEVCY-----AKTLANGSSMIVPKOKRLSASVEKEEL--CVKYEQWSESD 112
Db 113 tsttqlpcsatpftfgdllaangqg-----qgrritsvqprpqgqewlkmfqsgpe 167
QY 113 QVERVEHLISQMKYCHGHTNSYLRKMLQDRDFTLALPARGIDHAEMLISYLDKSLCAA 172
Db 168 KLAIdelIdscptqkmmqylepqfgrdfisllp---kelalyvlstlepkdlqga 223
QY 173 ELVCKEYRVTSDDMLMKKRLIERMVRTDSLWRGLAE-----RRMGQYLFRKNRPDGNAP 227
Db 224 agcrryrrllaednlvrekce-----egldpehlhkrk-----vlpgfihsp 265
QY 228 PMSFYRALYRKIIQDITFESNMRCGRHSIQRIHCRSETSKGVYCLQYDQKIVSGLRDN 287
Db 270 wksey-----lrg--hrldtnwrrgelksrkv-lkghdhdvltclqfcgqnlvsgsdn 320
QY 288 TIRIMDKNTLECKRILNGHTSGVYCLQYDERVITITSSDSTVRVWDMVTGEMLNTLIHNC 347
Db 321 CLKWSAVTGKCLTLVGHGCGVWSSGMDHLLISGSDTLIKWMAELGECILTYGHT 380
QY 348 EAVLHLRFNNGMNVYCSKDRSIAVWDMASPTDILRLRVLGHRRAVNVDDDKIVSAS 407
Db 381 styrcmhlhekryvsgsdcltlrwdletqcl---hvlmgthvaavvcygdgrivvsga 437
QY 408 GDRTIKWNMTSTCEPVRFLNGHKRGICLQYDRDLVWSSGSDNTIRLMDIECGACRLVLE 467
Db 438 ydfmrvkwvdpetclhtlqghtrvyslqfdqihvsvgsldcslrwdvetsgncilht 497
QY 468 GHEELVRCIRFDMKRIYSGVADGRIKWDVLAALDPRAPAGTLCRLTV---EHSGRVFR 524
Db 498 ghqsltsymelkanilysgnadsclvkiwdlktg-----qclqtlgqpnkqhsvatc 548
QY 525 LOFDEQIVSSSHDDTLIMD 545
Db 549 lqfknltvltssddgtvklwd 569

RESULT 12

B59201 B59201 standard; protein; 626 AA.
AC B59201;
DT 23-MAR-2001 (first entry)
DE Protein encoded by construct gmyc-N-sel-10.
XX Sel-10; human; Alzheimer's disease; Abeta.
XX Synthetic.
XX WO200075328-A1.
XX 14-DEC-2000.
XX 23-MAY-2000; 2000WO-US09814.
XX 09-JUN-1999; 99US-0328877.
XX (PHAA) PHARMACIA & UPJOHN CO.
XX Pauley AM, Gurney ME, Li J;
XX WPI: 2001-102404/11.
XX New human sel-10 polypeptides and their encoding polynucleotides,
XX useful for raising antibodies for detecting sel-10 polypeptide

Db 431 nadstvkikwdiktgy-----gcqlgtlqgpnkhgsavtclqfknkvitssdgdgtvkl 481
Qy 544 WD 545
Db 482 wd 483

RESULT 14
Y22468
ID Y22468 standard; Protein; 540 AA.
XX
AC Y22468;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human mammary sel-10 protein sequence.
XX
KW Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW mammary gland; therapy.
XX
OS Homo sapiens.
XX
PN MO9932623-A1.
XX
PD 01-JUL-1999.
XX
PF 17-DEC-1998; 98WO-US26820.
XX
PR 19-DEC-1997; 97US-0068243.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Gurney ME, Li J, Pauley AM;
XX
DR WPI; 1999-458026/38.
XX
DR N-PSDB; X99702.
XX
PF New isolated human sel-10 polypeptides
XX
PS Claim 24; Page 66-69; 91pp; English.
XX
XX
XX This sequence represents a human sel-10 protein of the invention. This
CC sequence is specifically a human mammary sel-10 protein. The polypeptides
CC can be used to alter presenilin function. Compounds which inhibit either
CC the expression or the activity of the human sel-10 polypeptides may
CC reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC therefore may be useful for the prevention or treatment of Alzheimer's
CC disease.
XX
XX
SQ Sequence 540 AA;

Query Match 21.1%; Score 640; DB 20; Length 540;
Best Local Similarity 30.8%; Pred. No. 1.3e-53;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

Qy 14 KFMNSEREDCNNGEPPKRIIEKNSLRQYNSCARLCLNCFVLASTAMKTEVCVAKT 73
Db 4 klhgservrslslgkpkckv-----seytstgl-----vpsca-----tpitfgdl 45
Qy 74 KLANCTSSNIYVKKOKKLASYTEKEKL--CYKYEQMSBDSQVEFVEHLISOMCHYQHGH 131
Db 46 raanqgq-----qgrritlsvqptglqewlkmfqswepekllaldeldisceptqvk 100
Qy 132 INSYLKPMLOPFTALPARGIDHIAENILSYLDAKSLCAAEELVCKEWRVRYTSDGMLMKK 191
Db 101 mmqvlepfqfqlfslp---keialyvlstlepkdlldaaqctrryrliaednllwre 156
Qy 192 LIERVVRTDSLMRGLAE-----RGMGOYLFEKNKPPDGNAPNSFYRLYKIIIDIEFI 246
Db 157 kcke-----egldelphlkrirk-----vlpkgflhspxksay-----lrg-hri 194
Qy 247 ESNMHCGRHSIORICRSETSKGYCLOYDOKIVSGLRDNTIKIMDKNTLECKRILTGH 306

Db 195 dtuwrigelkspkv-lkghddhvtclclqfagnrlyvsaddnllkwsavtgkclrtlvgh 253
Qy 307 TGSVLCLOQYDERVITITGSSDSVTWVDVTGEMLTLLHHCFAVLAHLRFNNOMVTCGSD 366
Db 254 tggvssqmdnlllsgatdrllkxwaaetgecintllyghstvcvcmllhckrvvsgsrd 313
Qy 367 RSIAYMDASPTDITLRRVLGHRAAVNVDFDKYIYASASGDRITIKWNTSTCEFEVRL 426
Db 314 atlrwvdieltgcl---hvlmgvaavrcvgydgrirrvsagydfnkvkwpdetcclhrl 370
Qy 427 NGHKRGIACTQYRDRLVYSSGSDNTIRLMDICGACLRVBESEELVACIRPDNKRIYSG 486
Db 371 gghlnrvyslqfdgihvvsqsltslrwvdeitgncintllyghsiltsgmelkdnllvsq 430
Qy 487 AYDGKIKYWDLYAALDPAPRAGTLCRLTV---EHSGRVFRLOPFEFOIVSSHDDTILI 543
Db 431 nadstvkikwdiktgy-----gcqlgtlqgpnkhgsavtclqfknkvitssdgdgtvkl 481
Qy 544 WD 545
Db 482 wd 483

RESULT 15
B59197
ID B59197 standard; protein; 540 AA.
XX
AC B59197;
XX
DT 23-MAR-2001 (first entry)
XX
DE Human hippocampal sel-10-5 protein.
XX
KW Sel-10; human; Alzheimer's disease; Abeta.
XX
OS Homo sapiens.
XX
PN WO200075328-A1.
XX
PD 14-DEC-2000.
XX
PF 23-MAY-2000; 2000WO-US09814.
XX
PR 09-JUN-1999; 99US-0328877.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Pauley AM, Gurney ME, Li J;
XX
DR WPI; 2001-102404/11.
XX
XX
XX New human sel-10 polypeptides and their encoding polynucleotides,
PT useful for raising antibodies for detecting sel-10 polypeptide
PT expression and as drug targets in the treatment of Alzheimer's disease
PT
PS Claim 1; Page 75-78; 116pp; English.
XX
XX The present invention relates to human sel-10. The sel-10 proteins of
CC the invention are useful for raising monoclonal or polyclonal
CC antibodies useful in diagnostic assays for detecting sel-10
CC polypeptide expression. The sel-10 polypeptides are also useful as drug
CC targets for decreasing antibody levels in the treatment of Alzheimer's
CC disease. It is also useful for identifying agents capable of
CC altering the production level of Abeta. The polynucleotides are useful
CC for developing assays for identifying agents capable of interfering
CC with the biological pathways that lead to Alzheimer's disease.
XX
XX
SQ Sequence 540 AA;

Query Match 21.1%; Score 640; DB 22; Length 540;

Best Local Similarity 30.8%; Pred. No. 1.3e-53;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```
OY 14 KFMNSSEEDCINNGEPKRIIEPKNSLRQTYNSCARCLINQETVCLASTAMKTENCVAKT 73
Db 4 kldhgsvevstslgkpkckv-----seytstcgl-----vpcsa-----tpftfgdl 45
OY 74 KLANGTSSMIVPKORRLSAYEKREL--CVKYPQMSSESDQVEFEVHLISQMCYOHGH 131
Db 46 raanggg-----qgrritsvqptglqgwlkmfsgswgpekllaladelidsceptqvk 100
OY 132 INSYLKPMLOPDTALPARGLDHAENILSYLDAKSLCAELVCKEMRYRTSDGEMWKK 191
Db 101 mmqvleqpqfgrdfisllp-----kelalylsflepkdllqaagctcrywrlaeadhllwre 156
OY 192 LIERAVRTDSLMKGLAE-----RGMGOYLFRKNRPDGNAPNSFYRALYPKIIDIETI 246
Db 157 kcke-----egldephllkrk-----vfkpgflhspsksay-----lrg--hrl 194
OY 247 ESNMRCGRHSLORIHCRSETSKGVYCLQYDDQKIYSGLRDNTIKIMDKNTLECKRILTGH 306
Db 195 dcnwrrtgelkspkv--lkghddhvtclqfcgnrtivsgsdnclikwsavtgcrltlvgh 253
OY 307 TGSVLCLOYDERVITIGSSDSTFVRWVDVNTGEMLTLIHCEAVLHLRFNNGMNVTCSDK 366
Db 254 tgvvssqmrndhllsgetcltkvnaetgecithlyghtstvtzcmhlhekrvvsgrd 313
OY 367 RSLAVWDMASPTDITLRRYLVGHRRAVNVDFDDKYIVASGDRTIKVMNTSTCEFVRTL 426
Db 314 atlrwddietcgcl---hvlmgvaavtcvgdygrrvvsgaydfmfvkvwdepelctclhll 370
OY 427 NGHKRGIACTQYRDRLVVSSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDRNRIVSG 486
Db 371 qghtrrvyslqfdgihvsgsldstlrwddetgncihcltghsltsgmekdnhlvsq 430
OY 487 AYDGKIKVMDLYAALDPRAPAGTLCIRTLV---EHSGRVFRLOQDFEQIVSSSHDDTILI 543
Db 431 nadstvkikwdikty-----gclqtlgspnkhsavtclqfnknfvltssddgtvkl 481
OY 544 WD 545
Db 482 wd 483
```

Search completed: June 7, 2001, 10:35:31
Job time: 30 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:35:01 ; Search time 13.8 Seconds
(without alignments)
792.100 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034 1 MDPAEVLDKALKFNNSSE.....PAAQADPPSPSTYYISR 569

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_Aa.*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2582.5	85.1	517	1 US-08-190-802A-30	Sequence 30, Appl.
2	520	17.1	587	3 US-08-899-578-2	Sequence 2, Appl.
3	399	13.2	779	1 US-08-190-802A-32	Sequence 32, Appl.
4	354	11.7	409	2 US-08-283-917-3	Sequence 3, Appl.
5	354	11.7	409	2 US-08-961-716-3	Sequence 3, Appl.
6	354	11.7	410	2 US-08-283-917-9	Sequence 9, Appl.
7	354	11.7	410	2 US-08-961-716-9	Sequence 9, Appl.
8	339.5	11.2	409	1 US-08-190-802A-51	Sequence 51, Appl.
9	321.5	10.6	514	1 US-08-190-802A-66	Sequence 66, Appl.
10	318	10.5	422	1 US-08-190-802A-52	Sequence 52, Appl.
11	313.5	10.3	209	3 US-08-899-578-6	Sequence 6, Appl.
12	306	10.1	704	1 US-08-188-582-5	Sequence 5, Appl.
13	306	10.1	704	1 US-08-646-715-5	Sequence 5, Appl.
14	305	10.1	704	1 US-08-190-802A-62	Sequence 62, Appl.
15	305	10.1	704	2 US-08-308-818-3	Sequence 3, Appl.
16	298.5	9.8	704	1 US-08-188-582-18	Sequence 18, Appl.
17	298.5	9.8	704	1 US-08-646-715-18	Sequence 18, Appl.
18	285.5	9.4	798	1 US-08-190-802A-64	Sequence 64, Appl.
19	285.5	9.4	798	1 US-08-190-802A-68	Sequence 68, Appl.
20	285.5	9.4	798	3 US-08-308-818-2	Sequence 2, Appl.
21	283.5	9.3	713	3 US-08-899-578-7	Sequence 7, Appl.
22	283	9.3	713	1 US-08-190-802A-63	Sequence 63, Appl.
23	263	8.7	318	1 US-08-190-802A-33	Sequence 33, Appl.
24	256	8.4	906	1 US-08-190-802A-31	Sequence 31, Appl.
25	245.5	8.1	317	1 US-08-190-802A-27	Sequence 27, Appl.
26	245.5	8.1	317	1 US-08-190-802A-41	Sequence 41, Appl.
27	245.5	8.1	317	1 US-08-190-802A-47	Sequence 47, Appl.

28	239	7.9	2627	2 US-08-751-189-3	Sequence 3, Appl.
29	239	7.9	2627	2 US-09-060-836-3	Sequence 3, Appl.
30	239	7.9	2627	4 US-09-184-445-3	Sequence 38, Appl.
31	236	7.8	340	1 US-08-190-802A-38	Sequence 42, Appl.
32	235	7.7	340	1 US-08-190-802A-42	Sequence 1, Appl.
33	233	7.7	395	3 US-09-032-372-1	Sequence 39, Appl.
34	232.5	7.7	326	1 US-08-190-802A-39	Sequence 4, Appl.
35	226	7.4	2629	2 US-08-751-189-4	Sequence 4, Appl.
36	226	7.4	2629	2 US-09-060-836-4	Sequence 4, Appl.
37	226	7.4	2629	4 US-09-184-445-4	Sequence 1, Appl.
38	222	7.3	305	3 US-08-965-600-1	Sequence 40, Appl.
39	215	7.1	340	1 US-08-190-802A-40	Sequence 45, Appl.
40	208.5	6.9	341	1 US-08-190-802A-45	Sequence 1, Appl.
41	203	6.7	350	2 US-08-828-922-1	Sequence 65, Appl.
42	196.5	6.5	439	1 US-08-190-802A-65	Sequence 2, Appl.
43	194.5	6.4	325	1 US-07-626-589-2	Sequence 57, Appl.
44	194.5	6.4	325	1 US-08-190-802A-57	Sequence 1, Appl.
45	194.5	6.4	325	1 US-08-323-444A-1	

ALIGNMENTS

RESULT 1
US-08-190-802A-30
Sequence 30, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Theoreof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190, 802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fablian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-190-802A-30

Query Match 85.1% Score 2582.5; DB 1; Length 517;
Best Local Similarity 91.4%; Pred. No. 4,7e-266;
Matches 487; Conservative 8; Indels 31; Gaps 2;

```

OY 18 S SEREDCNNGEPPKRIIPEKNSLROFYNSCARLNOETVCLASTAMKTECNVAKTKLAN 77
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13 ASEREDCNNGEPPKRIIPEKNTKLNQ-----TKLAN 42

OY 78 GTSSMIVPKQKRLSASAEKKEKELCVKFEQWSESDOVEFEHLISOMCHQOHGINSYK 137
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 43 GTSSMIVPKQKRLSASAEKKEKELCVKFEQWSESDOVEFEHLISOMCHQOHGINSYK 102

OY 138 PMLORDFTALPARGLDHIENILSYDAKSLCAELVCKEYRTVSDGLMKKLEIRMY 197
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103 PMLORDFTALPARGLDHIENILSYDAKSLCAELVCKEYRTVSDGLMKKLEIRMY 162

OY 198 RTDSLWGLAERRGQGYLFENKPPDGNAPNSFYRALYKIIODIETIESNMRGHS 257
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 163 RTDSLWGLAERRGQGYLFENKPPDGNAPNSFYRALYKIIODIETIESNMRGHS 222

OY 258 QRHCSESTSGYVCLQYDOQKIVSGLRDNTIKIMDKNTIECKRIILGHTGSVLCLOYDE 317
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 QRHCSESTSGYVCLQYDOQKIVSGLRDNTIKIMDKNTIECKRIILGHTGSVLCLOYDE 282

OY 318 RVITGSSDSTVRVMDVNTGEMLTILHHC EAVLHLEFNNGMVTC SKDRSIAVMDKASP 377
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 RVITG--SDSTVRVMDVNTGEMLTILHHC EAVLHLEFNNGMVTC SKDRSIAVMDKASA 341

OY 378 TDITLRLVYVGHRAAVNVDFDCKYIVSASGDRITIKVWNTSTCEVFTLNGHKGIAQLQ 437
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 342 TDITLRLVYVGHRAAVNVDFDCKYIVSASGDRITIKVWNTSTCEVFTLNGHKGIAQLQ 401

OY 438 YRDLRYVSGSSDNTIRLMDIECGACLVLESGHELVLCIRFDNKRIVSGAYDGKIKYMDL 497
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 402 YRDLRYVSGSSDNTIRLMDIECGACLVLESGHELVLCIRFDNKRIVSGAYDGKIKYMDL 461

OY 498 VAALDPRAPACTLCRTLVESHGVRFLQDFEFOIVSSSHDDITLLMDPLNDP 550
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 462 VAALDPRAPACTLCRTLVESHGVRFLQDFEFOIVSSSHDDITLLMDPLNDP 514

```

```

RESULT 2
US-08-899-578-2
: Sequence 2, Application US/08899578
: Patent No. 6087153
:
: GENERAL INFORMATION:
: APPLICANT: Greenwald, Iva
: APPLICANT: Hubbard, E. Jane
: TITLE OF INVENTION: SEL-10 AND USES THEREOF
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/899,578
: FILING DATE: 24-JUL-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AMC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 278-0400
: TELEFAX: (212) 278-0525
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 587 amino acids
: TYPE: amino acid

```

```

: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-899-578-2

```

```

Query Match 17.1%, Score 520; DB 3; Length 587;
Best Local Similarity 28.8%, Pred. No. 2e-46;
Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

```

```

OY 72 KTKLNGTSSMT-----VPKOR--LSASYEKEL-----CVKFPQWSESPOVE 115
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 35 ESSYSGSSSSYNADKLSSRPLOKHLDLSAPSPSNDLNRVEHLALFRLDLSAARDMD 94
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 116 FVEHLISQWCHYOHGINSYK PMLORDFTALPARGLDHIENILSYDAKSLCAELV 175
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 AFTRLQESNMNTINQRLAIIIEPHQRFLOSLPV---ELGMLHLNLGYDILKAAOV 150
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 176 CEWTRVTSQGLMKKL--IERMV-----RTDSLWGLAERRGQGYLFENKPPDGNAP 227
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 SKMWKLISEIDKIMKSLGVEEFKHPDPTDRYTGAMQCTATAAG-----VTIPDHQ 203
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 228 PN-SFYRALYKIIODI-----ETIESNMRGHSIQRIHCRSETSKGYCL 273
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 CDLWVHREFLKLOKFGDIFERRADKSRYLRADKIENNMNANPIMGSAY--LRGHEHVI 262
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 274 QYDDQKIVSGLRDNTIKIMDKNTIECKRIILGHTGSVLCLOYDE--RVITGSSDSTVR 331
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 QIHDVLYVSGSDNTLKWCIIDKGEVMTLVGHTGCVTSGISQCGRIYVSGSDTRIVKV 322
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 332 WDVNTGEMLTILHHC EAVLHLEFNNGMVTC SKDRSIAVMDASPDITLRLVYVGHRA 391
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 WSTVDGSLHTLQGHSTVRCAMAGSILVIGSRDPTTLRVMDVESGRILA---TLHGHA 379
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 392 AVNVVDFDCKYIVSASGDRITIKVWNTSTCEVFTLNGHKGIAQLQYDR--LVVSGSSD 449
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 AVRCVQDFGTIVSGCYDFTYAIMAHGRCIRITLTGNANNVYSLFSESESYICSGSLD 439
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 450 NTIRLMDI---ECGACLVLESGHELVLCIRFDNKRIVSGAYDGKIKYMDLVAALDPRAP 506
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 TSIKRVWDFRPGQECVALLQGHSTLSGMOQLRGMTILVSCANDSHVRWMDI-----H 491
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 507 AGTLCRTLVESHGVRFLQ--FDEFOIVSSSHDDITLLMD 545
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 492 EGT-CVHMLSGHRSATISLQWFGRMVATSSDDGTIVKLM 530
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 3
US-08-190-802A-32
: Sequence 32, Application US/08190802A
: Patent No. 5519003
:
: GENERAL INFORMATION:
: APPLICANT: Mochly-Rosen, Daria
: APPLICANT: Ron, Dorit
: TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
: NUMBER OF SEQUENCES: 265
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: P.O. Box 60850
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306-0850
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/190,802A
: FILING DATE: 01-FEB-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:

```


NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33, 875
 REFERENCE/DOCKET NUMBER: 8600-0139
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 779 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: Protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CD4 / CDC20 protein, Fig. 15
 US-08-190-802A-32

Query Match 13.2%; Score 399; DB 1; Length 779;
 Best Local Similarity 24.3%; Pred. No. 2.4e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TYCLASTAKTENCYAKTKLNGTSMIYPRKQKLSASTEKK-----ELCVKTFEOW 108
 DB 194 TPLAKTKTKTINN-----NNNIADLIESKDSIISPELSEISAINNNLPHAYFK-- 244
 QY 109 SSSDOVEFEHLISOMCHYOHGINSYKPMLODPETALPARGLDHIAENITSLYDAKS 168
 DB 245 -----NLFLRYVANDSELSDLGLIKDNKRDITSLP---EISLKIFNVLQFED 294
 QY 169 LCAAEVCEKMYR-VTSDMLMKLI--ERAVRTDSIMRGLAERRNGOYLFRKNRPPDN 225
 DB 295 IINSLGVSQNMKIRKSTLKKLISENFV-----SPKGF 331
 QY 226 APPNSFYRLYKTIOD-----IET--IESNMRCGRSLDRIHCRSETSKGVCLQYD 276
 DB 332 NSLNKLKSOYKPKLSQODRLRISFLENFILLKNWYNPKFVPORTLRLGHTSVITCLOPE 391
 QY 277 DQKIVSGLDNTIKIMDKNTLECKRILTGHTGSVYLCLOYDE--RVIIIGSSDSSTVRVMDVN 335
 DB 392 DMYVTGTGADDKMIRYDSINKKFLQLSGHDGVAALKRAHGILVSGSTDTKTRVMDIK 451
 QY 336 TGEMLNTLIHCEAVLHLRFNNGMMVTCSDRSIAVMDASPTDITLRLVGLHRAAVNV 395
 DB 452 KG-----CCT-----HVFEGHNSIVRC 468
 QY 396 VDFDD-----KYIVSASGRTIKVNT-----STCE-----FRTLNG 428
 DB 469 LDIVYKNIKIKYIVTGSRDNTLHWKLPRESSVPDHGEHDYPLVFTPEENPYFVGVLNG 528
 QY 429 HRRGIACIQYRDRLVYSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN--KRIYSG 486
 DB 529 HMASTRVYSGHNIYVSSGYDNTLIYMDVAKMKCLYILSGHIDRIYSTIYDHEKRCISA 588
 QY 487 AYDGKIKWDL-----VAALDPRAPAGTL--CLRTLVHSGRVERFLQDEFQIVSS 535
 DB 589 SMDTIRIDNLEINWNGECYSATNSASPCAKILGAMTYLQGHATVGLRLSDRELVSA 648
 QY 536 SHDDTILIMDFLNDPAAQAEPRPSRTITY 566
 DB 649 AADGSIRGMD-AND-----YSRKFSTY 668

RESULT 4
 US-08-283-917-3
 ; Sequence 3, Application US/08283917
 ; Patent No. 5849557
 ; GENERAL INFORMATION:
 ; APPLICANT: ADACHI, HIDEKI
 ; APPLICANT: TSUJIMOTO, MASAFUMI
 ; APPLICANT: INOUE, KEIJO
 ; APPLICANT: ARAI, HIROYUKI

TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
 TITLE OF INVENTION: AND GENE THEREOF
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
 ADDRESS: NEUSTADT, P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/283,917
 FILING DATE: 03-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 209943/1993
 FILING DATE: 03-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, No. 5849557/man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2292-030-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 409 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Bos taurus
 US-08-283-917-3

Query Match 11.7%; Score 354; DB 2; Length 409;
 Best Local Similarity 29.2%; Pred. No. 5e-29;
 Matches 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;

QY 280 IYSGLDNTIKIMDKNTLECKRILTGHTGSVYLCLOYDE--RVIIIGSSDSSTVRVMDVNTG 337
 DB 122 MVSASEDAIKWMDYETGDFERTLGHGHSYODISFDHSGKLASCADMTIKLMDPQGF 181
 QY 338 EMLNTLIHCEAVLHLRF--NNGMMVTCSDRSIAVMDASPTDITLRLVGLHRAAVNV 395
 DB 182 ECFIRMHGHNVSSVALIMPNGDHIVASRDRTIKMVEYQYGYCY---KTFGHHKVRM 238
 QY 396 V--DEDDKYIVSASGRTIKVNTSTCEFVRLNCHKRGIACTQYRDR----- 441
 DB 229 VAPNDGTLIASCNDQYRVVAVATKECKALREHVEHVECIINAPSSYSISEATGS 298
 QY 442 -----LVYSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN--KRIYSGAYD 489
 DB 299 ETKSKGPGPFLSSGRDKTIKMPVSTGMCIMTLVGHDMVWVGLFHSFGGFFILSCADD 358
 QY 490 GKIKWMDVLAALDPRAPAGTLCRTLVHSGRVERFLQDEF--QIVSSSHDDTILIMD 545
 DB 359 KILRWMDY-----KNKRCMKTLNNAHEHVTSLDPRKTAIPYVYVTSVQTVAYWE 407

RESULT 5
 US-08-961-716-3
 ; Sequence 3, Application US/08961716
 ; Patent No. 5880272
 ; GENERAL INFORMATION:

Db 359 KTLRWMDY-----NKKRCKMTLNAHEHFTVSLDFHKTAPYVVTGSDQTVKWE 407

RESULT 9

US-08-190-802A-66
Sequence 66, Application US/08190802A
Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190, 802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fadian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 514 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: YCWX PROTEIN, Fig. 49

US-08-190-802A-66

Query Match

Best Local Similarity 10.6%; Score 321.5; DB 1; Length 514;

Matches 88; Conservative 44; Mismatches 97; Indels 59; Gaps 9;

Db 236 SIQRHCSESESKYCYCLOYDQKIVSGIRDTIKIMKNTLECKRIITGHTGSLCLOY 315

Db 236 SWEPIHLVPGSK-----PRIASSKDGITIKIMPTVSCQYTSNGHTNSVCWK 286

Db 316 -DERVITGSSSTYRVMDVT-GEMLNTLIHCEAVLHLREN----- 356

Db 287 GGGLLYSGSHDRYRVMDINSQRCINILKSHAMVNHLSITDYALRIGAFDHTGKKP 346

Db 357 -----NG---MMVTCSDRSIAVMDMASPTDITLRRVYVGHRAA 392

Db 347 STPEBAQKALENYEKICKKNGNSEMMVFTASDYTMFLMPLKSTKPLAR--MTGHOKL 404

Db 393 VNVVDF--DDKYIVASASDRITIKYNTSTCEVRLNKHKGCIACLOYRD--RLVSSSS 448

Db 405 VNHVAFSPDGRIRYVASPDNSIKIMDGDGKFIPTFRGHIAVYQVAVSSDCRIILVSCSK 464

Db 449 DNTIRLMDIEGACLRVLEG--HEELVRCIRFQDNKRIVSGAVDGKIKVW 495

Db 465 DTTIKVMDYVTRKRLSVLDLPGIKTKILYVMSVDGKRVCSGGGDKMVRW 512

RESULT 10

US-08-190-802A-52
Sequence 52, Application US/08190802A
Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190, 802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fadian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 52:

SEQUENCE CHARACTERISTICS:

LENGTH: 422 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: MD6, Fig. 35

US-08-190-802A-52

Query Match

Best Local Similarity 10.5%; Score 318; DB 1; Length 422;

Matches 113; Conservative 69; Mismatches 161; Indels 120; Gaps 16;

Db 103 KYFOWSESDQVER-----VEHLISOMCHYGHINSYKPMLODFITALPAR 151

Db 4 KDFETWLDNIVTEPLSLMDLOKNETLDHLISLGAVALRHLSNNLETLKRDPLKPL 62

Db 152 GLDIAENILSYLDAKSICAEVLCKEYRVTSQGMILKLIEMRVRTDSIMGLAERG 211

Db 63 ---ELSFYLLKLPDQFLITLCCIVSKQNNKVIS-----ACTEWAQACNLG 106

Db 212 WGVLFKMKPPDGNAPRPSFRLALYPKTIODIETIESNWRGRHSLSRHRORSSTSGVY 271

Db 107 W-----QIDDSVQDSLHWKKVYLKAILRMKQLED-----HEAFETSS--- 143

Db 272 CLOYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLCLOYDERVITGSSDSTVRV 331

Db 144 -----LHISARVYALYYKKDGLCTGSDLSAKL 172

Db 332 WDVNTGEMLNTL-IHCEAVLHLRFNNGMAYTCSKDSIAVMDMASPTDITLRRVYVGH 390

Db 173 WDVSTGCGVYVGIQHTCAV---KFDEOKIVTGSFDMTVACMEWSSGARQOHFR---GHT 226

Db 391 AAVNVDFDCK--YIVASGDRITIKYVNTSTCEVRLNKHKGCI-----ACLO 437

Db 227 GAVESVYDELDILVSSADPAVKVWALSAGCLNTLGHTEWTKVYLQCKVSLH 286
QY 438 YRDLVYSSSDNTIRLW---DIEGACLRVLEGHLEEVRCIR---EDNKRIVS--- 485
Db 287 SPEDYILLSADKREIKIMPIGREINC-KCLKTLVSSEDSICLOPRLHDGKIVYSSAL 345
QY 486 GAYDGKIKWDLVLAALDPAPACITLCLRTLVHSGHVFRLQD 528
Db 346 GLYWDFAASYDILRLVKTPEVANLALL-----GFGDVALLFD 383

RESULT 11
US-08-899-578-6
; Sequence 6, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,578
FILING DATE: 24-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/53200/JPM/AKC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 278-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-899-578-6

Query Match 10.3%; Score 313.5; DB 3; Length 209;
Best Local Similarity 28.8%; Pred. No. 3.3e-25;
Matches 81; Conservative 37; Mismatches 80; Indels 83; Gaps 10;

QY 270 VYCIQYDDQKIYSGLRDNTIKIMDKNTLECKRIILGHGTSVLCLOYDE--RYIITGSSDS 327
Db 7 ITCOMQHDVLYVTGSDNTLKW-----CGHTGVTSTQISOGRIYVSGSTDR 55
QY 328 TVRVMDVNTGEMINTLIHCEAVLHLRFNNGMAMVTCSKDRSLAVMDASPTDITLRLVLY 387
Db 56 TVKVMGHSITV-----RCMAMA-----GSTLVYGSRTTILRVMD----- 90
QY 388 GIRAANVYDFPDKIYVSGSRTIKVNTSTCEVRYTLNGHKRGIALQYDR--LVVS 445
Db 91 GHAAARCVQFGTIVYSGGYDFYKIV-----NGHNRRYSLTFESERSIVCS 139
QY 446 GSSDNTIRLMDCEGACLVLEGHLEEVRCIRPDNKRIVSGAYDGKIKWDLVLAALDPA 505
Db 140 GSLDTSIRVMD-----GHTSLTSGMQLRGNTILVSCNADSHVRVMD----- 179

QY 506 PAGTLCRTLVHSGHVFRLQ-FDEQIVSSSHDITILWD 545
Db 180 -----GHRSAITSIQWFGRMVATSSDDGTVKLMD 209

RESULT 12
US-08-188-582-5
; Sequence 5, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dylact, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; NUCLEIC ACIDS ENCODING TATS AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RNO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-188-582-5

Query Match 10.1%; Score 306; DB 1; Length 704;
Best Local Similarity 22.8%; Pred. No. 1.6e-23;
Matches 122; Conservative 90; Mismatches 192; Indels 132; Gaps 22;

QY 95 EKEKELCVYFEQWSESDOVERVEHLISQM---HYQGHINSLY--KP--MLQRFIT 146
Db 160 EKAKETIEKY-----KCDLDGYYIEGLFNLLLSKPEELLENDLV 200
QY 147 ALP-----ARGLDHIAENITSLYDAKSLCAELVCKEYRYVSDGLMKLIERVR 198
Db 201 AEWQDKFYIRMSRDSLSLKRHIQ--DRQEVYADIVSKYLFHFDYEGARKKL--QCYA 256
QY 199 TDSLNGIAGRRGQCYLF-----KNRP----- 221
Db 257 TAGSHLGEAKRODNKRVRYYGLLKEVDFTLTTPAPAPEEDDDPDAPRPKKKPKPD 316
QY 222 -----PGNAPNPSFYALPKIITDIETIESNMRCGRHSLQRIHCS----- 265
Db 317 ILSKRSKSDPNAP--SIDRIPLPELKDSDKLK--LKALREASKRLALSQDLPSPAVFT 372

QY 266 ---TSKGYVCLQYDDOK--IVSGLRDNFTIKIMDKNTLECKRILNGHGSVLCLOYDERVI 320
 Db 373 VLNHOGVTCALISDSTMLACGFGDSVRIWSLTPAKRLTKDADS-----LREL 423
 QY 321 ITGSSDSTVRVMDVNTGEMTLNLHHCBAVLHLP--NNGMAYTCSKDRSIAVMDMASPT 378
 Db 424 DKESADINVRMLDDRSRGVTRSLMGHTGPVYRCAPAPENMNLSCSESDSTRLMSLTLWS 483
 QY 379 DITLRRVLGHRRAAVNVDF--DDKVIYASAGDRITIKYWNSTCEFPVTLNGHKGRIACL 436
 Db 484 CVVYTR---GHVYPWVDVRFAPHGYYFVSCSYDKTARLWATDSNOLAEVGHLSDDVCV 540
 QY 437 QY--RDRLVVSGSSDNTIRLMDIEGACIRYLEGHEELVRCIRFD--NKRIYSGAYDKI 492
 Db 541 QHPNSNYVATGSSDRYTRMLDMNTGOSVRLMTGHKGSVSLAFSACGRYIASSGVYDNI 600
 QY 493 KWMVLVAALDPRAPAGTLCRLTVHSGRVRLQF--DEFOIVSSSHDITLIMDF 546
 Db 601 IIMDL-----SNGSL-VTTLRLHRTSTVTTITFSRDGTVLAAGLDNNLTLIMDF 647

RESULT 13

US-08-646-715-5
 ; Sequence 5, Application US/08646715

Patent No. 5637686

GENERAL INFORMATION:

APPLICANT: Tjian, Robert

APPLICANT: Comati, Lucio

APPLICANT: Dyniaet, Brian D.

APPLICANT: Hoey, Timothy

APPLICANT: Rupert, Siegfried

APPLICANT: Tanese, Naoko

APPLICANT: Wang, Edith

APPLICANT: Weitzierl, Robert O.J.

TITLE OF INVENTION: DATA-BINDING PROTEIN ASSOCIATED FACTORS,

TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/646/715

FILING DATE: 09-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/188,582

FILING DATE: 28-JAN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 704 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-646-715-5

Query Match 10.1%; Score 306; DB 1; Length 704;
 Best Local Similarity 22.8%; Pred. No. 1.6e-23;
 Matches 122; Conservative 90; Mismatches 192; Indels 132; Gaps 22;

QY 95 EKEKELCVYFEQWSESDQVEFEVHLISQMC---HYOGHINSYL--KP--MLQRDEIT 146
 Db 160 EKAKEFEIKY-----KCDLDGYTEIGLEFNLLLSKPELLENDLVV 200
 QY 147 ALP-----ARGDLHIAENLISYIDAKSLCAAEYCKEWMRYTSDGMKKLIERVYR 198
 Db 201 AMEQDFVIRMSDSHSLKRRHIQ--DRQEVVAIVSKYLFHFDYEGMARNKL--QCVV 256
 QY 199 TDEIAMEGLAERQWQYLF-----KNKP----- 221
 Db 257 TASHHLEAKRQDNKKRIVYGLLEKEVDPTLTTPAPAREEDDDPADRPKKKKPKPKD 316
 QY 222 -----PDGNAPNSFYALYPKITIDITETESNMRCGRHSIORGRSE----- 265
 Db 317 LSKKSKSDPNAP--SIDRIPLEPKSDSKLLK--LKALREASKRLALSKDQPSAVEYF 372
 QY 266 ---TSKGYVCLQYDDOK--IVSGLRDNFTIKIMDKNTLECKRILNGHGSVLCLOYDERVI 320
 Db 373 VLNHOGVTCALISDSTMLACGFGDSVRIWSLTPAKRLTKDADS-----LREL 423
 QY 321 ITGSSDSTVRVMDVNTGEMTLNLHHCBAVLHLP--NNGMAYTCSKDRSIAVMDMASPT 378
 Db 424 DKESADINVRMLDDRSRGVTRSLMGHTGPVYRCAPAPENMNLSCSESDSTRLMSLTLWS 483
 QY 379 DITLRRVLGHRRAAVNVDF--DDKVIYASAGDRITIKYWNSTCEFPVTLNGHKGRIACL 436
 Db 484 CVVYTR---GHVYPWVDVRFAPHGYYFVSCSYDKTARLWATDSNOLAEVGHLSDDVCV 540
 QY 437 QY--RDRLVVSGSSDNTIRLMDIEGACIRYLEGHEELVRCIRFD--NKRIYSGAYDKI 492
 Db 541 QHPNSNYVATGSSDRYTRMLDMNTGOSVRLMTGHKGSVSLAFSACGRYIASSGVYDNI 600
 QY 493 KWMVLVAALDPRAPAGTLCRLTVHSGRVRLQF--DEFOIVSSSHDITLIMDF 546
 Db 601 IIMDL-----SNGSL-VTTLRLHRTSTVTTITFSRDGTVLAAGLDNNLTLIMDF 647

RESULT 14

US-08-190-802A-62

; Sequence 62, Application US/08190802A

Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theoreof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Denlinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190,802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 62:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 704 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: TRANSCRIPTION FACTOR T11F, Fig. 45
 US-08-190-802A-62

Query Match 10.1%; Score 305; DB 1; Length 704;
 Best Local Similarity 22.8%; Pred. No. 2,1e-23;
 Matches 122; Conservative 89; Mismatches 193; Indels 132; Gaps 22;

95 EKEKELCVKFEQWSESDQVEFVEHLISQMC---HYQGHINSYL--KP--MLQRFIT 146
 160 EKAKFEIKY-----KCDLDGYIEGLFNLLLSKPELLENDLVV 200
 147 ALP-----ARGIDHTAENILSTLDKSLCAELVCKEYRYTSDGMLKKLIERVR 198
 201 AMEDKFEVIRMSRDSHSLFKRHIO--DRROEVVADIVSKYLHFDYEGMARNTL--QCVA 256
 199 TDSLWRLGLAERGMQGYLF-----KNRP----- 221
 257 TAGSHLGEAKKQDNKMRVYGLLEKVEVDFQTLTPAPAEEDDDPDADPRKKKPKKDP 316
 222 -----PDGNAPPNSEFYRALYPKTIIDITETESNMRCGRHSLOIRHCRSE----- 265
 317 LLSKRSKSDPNAP--SIDRIPLEIKSDKLLK--LKALREASKRLALSKQDLPsAVEYT 372
 266 ---TSKGYVCCQYDDQK--IYSGLRDNTIKITMDKNTLECKRLTGHGTSVLCLOYDERVI 320
 373 VLNSHOGVTCAEISDDSTMLACGFGDSSVRIWSTLPANVRLTKDADS-----LREL 423
 321 ITGSSDSTVRVWVNTGEMLTLLHCEAVLHLRF--NNGMMVTCGSKDRSIAMVDMASPT 378
 424 DKESADINVRMLDSDSGEVTSLMGHTGPRYRCAPAPENMLLSCSESDSTIRLMLTWS 483
 379 DITLRRVLVGHRAAVNVDF--DDKYIVSASGDRITKVMNTSTCEFTVTLNGHKGIACL 436
 484 CVVYTR--GHVYPMVDVRFAPHGYYFVSCSYDKTARLAMDSDNALRVFVGHLSDDVCV 540
 437 QY--RDRLVYSGSSDNTIRLMDIEGACLRVLEGEHELVRICIRPD--NKRIYSGAYDGKI 492
 541 QFHPSNTVATGSSDRTVRLMDNMTGQSVRLMTGKHGVSLSAFACGRYLAAGSVVDHNI 600
 493 KVMDLVALDPRAPAGTLCRLTVLHSGRVFRIQF--DEFQIVSSHDDTLIIMWF 546
 601 IIMDL-----SMGSL-VTTLRLHSTVTYTTITFSRDGYVLAAGLDNNLTIIMWF 647

RESULT 15
 US-08-308-818-3
 Sequence 3, Application US/08308818
 Patent No. 5847077
 GENERAL INFORMATION:
 APPLICANT: Green, Michael R
 APPLICANT: Reese, Joseph C
 TITLE OF INVENTION: A No. 5847077e1 Fungal Multisubunit Protein
 TITLE OF INVENTION: Complex Critical for Expression of Fungal Proteins
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESS: Darby & Darby PC
 STREET: 805 Third Avenue
 CITY: New York
 STATE: New York
 COUNTRY: US
 ZIP: 10022

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/308,818
 FILING DATE: 19-SEP-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Ludwig, S. Peter
 REGISTRATION NUMBER: 25,351
 REFERENCE/DOCKET NUMBER: 0342/0A404
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-52707700
 TELEFAX: 212-753-6237
 TELEX: 236687
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 704 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 FRAGMENT TYPE: N-terminal
 ORIGINAL SOURCE:
 ORGANISM: D. melanogaster
 IMMEDIATE SOURCE:
 CLONE: TAFIT-80
 US-08-308-818-3

Query Match 10.1%; Score 305; DB 2; Length 704;
 Best Local Similarity 22.8%; Pred. No. 2,1e-23;
 Matches 122; Conservative 89; Mismatches 193; Indels 132; Gaps 22;

95 EKEKELCVKFEQWSESDQVEFVEHLISQMC---HYQGHINSYL--KP--MLQRFIT 146
 160 EKAKFEIKY-----KCDLDGYIEGLFNLLLSKPELLENDLVV 200
 147 ALP-----ARGIDHTAENILSTLDKSLCAELVCKEYRYTSDGMLKKLIERVR 198
 201 AMEDKFEVIRMSRDSHSLFKRHIO--DRROEVVADIVSKYLHFDYEGMARNTL--QCVA 256
 199 TDSLWRLGLAERGMQGYLF-----KNRP----- 221
 257 TAGSHLGEAKKQDNKMRVYGLLEKVEVDFQTLTPAPAEEDDDPDADPRKKKPKKDP 316
 222 -----PDGNAPPNSEFYRALYPKTIIDITETESNMRCGRHSLOIRHCRSE----- 265
 317 LLSKRSKSDPNAP--SIDRIPLEIKSDKLLK--LKALREASKRLALSKQDLPsAVEYT 372
 266 ---TSKGYVCCQYDDQK--IYSGLRDNTIKITMDKNTLECKRLTGHGTSVLCLOYDERVI 320
 373 VLNSHOGVTCAEISDDSTMLACGFGDSSVRIWSTLPANVRLTKDADS-----LREL 423
 321 ITGSSDSTVRVWVNTGEMLTLLHCEAVLHLRF--NNGMMVTCGSKDRSIAMVDMASPT 378
 424 DKESADINVRMLDSDSGEVTSLMGHTGPRYRCAPAPENMLLSCSESDSTIRLMLTWS 483
 379 DITLRRVLVGHRAAVNVDF--DDKYIVSASGDRITKVMNTSTCEFTVTLNGHKGIACL 436
 484 CVVYTR--GHVYPMVDVRFAPHGYYFVSCSYDKTARLAMDSDNALRVFVGHLSDDVCV 540
 437 QY--RDRLVYSGSSDNTIRLMDIEGACLRVLEGEHELVRICIRPD--NKRIYSGAYDGKI 492
 541 QFHPSNTVATGSSDRTVRLMDNMTGQSVRLMTGKHGVSLSAFACGRYLAAGSVVDHNI 600
 493 KVMDLVALDPRAPAGTLCRLTVLHSGRVFRIQF--DEFQIVSSHDDTLIIMWF 546
 601 IIMDL-----SMGSL-VTTLRLHSTVTYTTITFSRDGYVLAAGLDNNLTIIMWF 647

Thu Jun 7 10:44:14 2001

Search completed: June 7, 2001, 10:36:14
Job time: 73 sec

us-09-415-795-4.ra1

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:35:01 ; Search time 21.72 Seconds

(without alignments)
1800.338 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034

Sequence: 1 MDRAEAVLQEKALKFMSSE.....PAAQAEPPSPSRITYISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2597	85.6	518	2	B48088
2	1635.5	53.9	701	2	T16607
3	690	22.7	506	2	T50211
4	590.5	19.5	605	2	T38932
5	545	18.0	640	2	S49332
6	531.5	17.5	650	2	T46660
7	520	17.1	579	2	T22703
8	519.5	17.1	267	2	S62507
9	455.5	15.0	1356	2	T18521
10	453	14.9	775	2	T45136
11	399	13.2	779	2	S56245
12	396.5	13.1	703	2	T43557
13	354	11.7	409	2	S36113
14	354	11.7	410	2	S48052
15	341	11.2	515	2	S19487
16	339.5	11.2	777	2	T41075
17	337	11.1	1146	2	A55532
18	332.5	11.0	317	2	T46032
19	318	10.5	422	2	A56440
20	313.5	10.3	376	2	T19266
21	312.5	10.3	714	2	S56893
22	312	10.3	876	2	T51507
23	307.5	10.1	1693	2	S76086
24	305	10.1	704	2	S33263
25	299.5	9.9	1194	2	T03188
26	297	9.8	502	2	T41148
27	292	9.6	659	2	S38108
28	290.5	9.6	614	2	S58305
29	287.5	9.5	327	2	S48839

30	285.5	9.4	798	2	S34023	TATA box-binding p
31	283	9.3	713	2	JN0133	WD-40 repeat regul
32	280	9.2	494	2	T19550	hypothetical prote
33	278.5	9.2	380	2	T40283	beta-transducin -
34	277	9.1	586	2	T38992	WD-40 repeat regul
35	276.5	9.1	501	2	T27513	hypothetical prote
36	274	9.0	2241	2	T16064	hypothetical prote
37	273	9.0	325	2	T09613	probable GTP-bind
38	272.5	9.0	1008	2	T32986	hypothetical prote
39	271.5	8.9	817	2	S51445	probable membrane
40	270	8.9	451	2	S65162	hypothetical prote
41	270	8.9	473	2	T33805	hypothetical prote
42	270	8.9	504	2	T50983	probable pleiotrop
43	270	8.9	651	2	T50289	WD repeat protein
44	269.5	8.9	643	2	T41454	transcription init
45	265.5	8.8	325	2	T23309	hypothetical prote

ALIGNMENTS

RESULT 1
B48088
beta-transducin repeat-containing protein - African clawed frog
N:Alternate names: beta-ttcp
C:Species: Xenopus laevis (African clawed frog)
C>Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000
C:Accession: B48088
R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.
Mol. Cell. Biol. 13, 4953-4966, 1993
A>Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase
A:Reference number: A48088; MUID:93330289
A:Accession: B48088
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-518 <SP>
A:Cross-References: GB:M96268; NID:9295542; PIDN:AAA02810.1; PID:9295543
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
C:Keywords: duplication
F:431-462/Domain: WD repeat homology <WD1>

Query Match	Best Local Similarity	Score	2597	DB 2	Length	518
Matches	488	Conservative	7	Mismatches	8	Indels
30	285.5	9.4	798	2	S34023	TATA box-binding p
31	283	9.3	713	2	JN0133	WD-40 repeat regul
32	280	9.2	494	2	T19550	hypothetical prote
33	278.5	9.2	380	2	T40283	beta-transducin -
34	277	9.1	586	2	T38992	WD-40 repeat regul
35	276.5	9.1	501	2	T27513	hypothetical prote
36	274	9.0	2241	2	T16064	hypothetical prote
37	273	9.0	325	2	T09613	probable GTP-bind
38	272.5	9.0	1008	2	T32986	hypothetical prote
39	271.5	8.9	817	2	S51445	probable membrane
40	270	8.9	451	2	S65162	hypothetical prote
41	270	8.9	473	2	T33805	hypothetical prote
42	270	8.9	504	2	T50983	probable pleiotrop
43	270	8.9	651	2	T50289	WD repeat protein
44	269.5	8.9	643	2	T41454	transcription init
45	265.5	8.8	325	2	T23309	hypothetical prote

QY	18	SSEREDCNNGEPPRKTIIPKNSLRQTYNSCARICLQETVCLASTAMKRENCVAKTKLAN	77
DB	13	ASEREDCNDEPPRKTIIPKNTLQ-----TKLAN	42
QY	78	GTSSMIVPRQKLSAYEKEKEICVYFEQWSESDQVEFEHLISQMGYOHGHINSYLK	137
DB	43	GTSSMIVPRQKLSAYEKEKEICVYFEQWSESDQVEFEHLISQMGYOHGHINSYLK	102
QY	138	PMLODFITLALPARGLDIAENILSYLDKSLCAELVCKEWMYRVTSDDMLMKLLIERNV	197
DB	103	PMLODFITLALPARGLDIAENILSYLDKSLCAELVCKEWMYRVTSDDMLMKLLIERNV	162
QY	198	PRDLSMRGLAERGMQYLFKNKPPDGKTPNSFYKALPKTIIDLETESMMRGGRSL	257
DB	163	PRDLSMRGLAERGMQYLFKNKPPDGKTPNSFYKALPKTIIDLETESMMRGGRSL	222
QY	258	QRTHCSERSKGVYCYQYDQKIVSGIRLNTIKIMDKNTLECKRIITGHTSVLCQYDE	317
DB	223	QRTHCSERSKGVYCYQYDQKIVSGIRLNTIKIMDKNTLECKRIITGHTSVLCQYDE	282
QY	318	RYITIGSSDSYRVMDVNTGEMLNTLIHCEAVLHLENNMGAVTCSKDRSLAVDMASP	377
DB	283	RYITIGSSDSYRVMDVNTGEMLNTLIHCEAVLHLENNMGAVTCSKDRSLAVDMASP	342
QY	378	TPITLRRVLYGHRVAVNVVDPDKYIVSASGRTIKYVNTSCFEPRTLNGHKGRIACIQ	437
DB	343	TPITLRRVLYGHRVAVNVVDPDKYIVSASGRTIKYVNTSCFEPRTLNGHKGRIACIQ	402

Oy 438 YRDLRVGSSSSNTIRLMDIEGACLRVLEGGHEELVPCIRPDNRKIRVSGAYDGKIKVMDL 497
 |||||||
 Db 403 YRDLRVGSSSSNTIRLMDIEGACLRVLEGGHEELVPCIRPDNRKIRVSGAYDGKIKVMDL 462
 |||||||
 Oy 498 VVALDPRAPAGTLCRLTVLVEHSGRVRFLQPFDEFOIVSSSHDDTILIMPLNDP 550
 |||||||
 Db 463 VVALDPRAPAGTLCRLTVLVEHSGRVRFLQPFDEFOIVSSSHDDTILIMPLNDP 515
 |||||||
 RESULT 2
 T16607
 hypothetical protein K10B2.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16607
 R:Miller, N.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of *C. elegans* cosmid K10B2.
 A:Reference number: Z18545
 A:Accession: T16607
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-701 <MIL>
 A:Cross-references: EMBL:U28730; NID:9860694; PID:9860695; PIDN:AAA68258.1; CESP:K10B2.1
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:K10B2.1
 A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 666/3

```

Query Match          53.9%: Score 1635.5; DB 2; Length 701;
Best Local Similarity 57.7%: Pred. No. 7,6e-119;
Matches 328; Conservative 69; Mismatches 116; Indels 55; Gaps 8;

Oy 31 KRIPEKNSLNQTYNSCARCLINDETVCLASTAMKTENCVAKTKLAN-----GTS 80
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 2 RRFREGKALTKRGHARGGSGIALTVGVST----IERGCF--TAVSNDFLEFSTFSFVF 55
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 81 SMIVKOR-----KLSASYEKKEKELCVKVFQWSESDQVEFEVHLSQMCHYOHGHIINSY 135
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 56 SFLPFSRMTQFLFSTRFSFSFEVL-----KWSHEQDLQDNKTVHRLSHYOLQKVNFE 110
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 136 LKPMLODFITALPARGLDHAENILSYLAKSLCAELVCKEMYRVTSIDGMLMKKLLIER 195
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 111 IRPMLQRFIFSLTPA---HLVELTIFVNSDSLKCEESTSMRCALARQGHMKKLTIEK 166
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 196 MYRDTSLWRGLAERGMGOYL-----FKNNPQGNAPNPSFYAL 235
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 167 NRSDSLWMMJSEKQWQKFLNISRDMSVRRICEKENYDVNIKRDKLDLILMHFYSKL 226
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 236 YPKIIDQDIETIESNNMRCGRHSIORIHCRSFETSKGYVCLQYQDOKIVSGLRDNTIKIMPN 295
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 227 YPKIIRDIHNDNNMKRGNYKMKTRINCQSENSKGYVCLQYQDDKIVSGLRDNTIKIMDK 286
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 296 TLECKRIITGHGVCLOXYDERVITITSSSDSYVWVWVNGEMLNTLIHCCEALHLRF 355
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 287 DYSCKRIISGHGVSVCLOYDNRVITISGSDPYVWVWVEGECKITLIHCCEALHLRF 346
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 356 NNGMNVTCSSKDRSLAVWDMASPTDITLRRVYLVGRHAANVVDVDDKYIVTSASGDTIKW 415
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 347 ANGLIVTCSKDRSLAVWDMVSPDITIRRVLYVGRHAANVVDVDDRYIVTSASGDTIKW 406
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 416 NTSTICEFRTLNGHKGKAGIACLOYDRILVYSSSDNTILMLIEGACALRYLGEHELVC 475
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 407 SMDTLEFRTLNGHKGKAGIACLOYDRILVYSSSDNTILMLIHSGVCALRYLGEHELVC 466
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 476 IRPNKRIYASGAYGKIVMDLVVALDPRAPAGTCLTLEHSGRVRLQFDEQIYSS 535
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 467 IRPEKRIYASAYGKIKVMDLQALDPRALSTELCISLVQHGVRVRLQFDDQIYSS 526
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Oy 536 SHDITILIMDFLNDPAAQEPFRSBRKT 563
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Db 527 SHDDTILWDFLADP-----PSGLPSST 549

RESULT 3

T50211

WD-repeat protein [imported] - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 02-Sep-2000

C:Accession: T50211

R:McDougal1, R.C.; Rajadream, M.A.; Barrel1, B.G.; Brown, S.; Murphy, L.; Jones, L.;

Submitted to the EMBL Data Library, January 2000

A:Reference number: Z25046

A:Accession: T50211

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-506 <MCD>

A:Cross-references: EMBL:ALJ36538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05

A:Experimental source: strain 972h(-); cosmid c30

C:Genetics:

A:Gene: SPAC296E.01; SPDB:SPAC30.05

A:Map position: 1

A:Introns: 43/1; 74/3

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

[illegible]

submitted to the EMBL Data Library, April 1997
 A: Reference number: Z31818

A; Reliance number: Z21818

A;Accession: T38932

A;Status: preliminary; translated from GB/EMBL/DBJ

A;molecule type: DN

A;Residues: 1-605 <BAD>

A: Cross-references: EMBL: Z94864: PTDN: CAB08168.1: GSPDB: GN00066: SPDB: SPAC57A10.05c

A: Experimental source: strain 972b-: coccolid 057A10
a, ci055 teleferences: EMBL:234804; PIDN:CAB06108.1;

A; Experimental

C;Genetics:

A;Gene: SPDB:SPAC

```
A;Map position: 1
```

Query Match	19.5%	Score 590.5	DB 2	Length 605
Best Local Similarity	29.0%	Pred. No. 6.7e-38		
Matches 156; Conservative	86	Mismatches 189	Indels 107	Gaps 12

QY	80	SSMIVPQPKRTKLSASYEKEKELCVAYFEQOMSE---	SDQVEFVEMHLSIQMCHYQGHHSINYL	136
Dd	46	SSM-----HNELSGISEKSRQREKVENMAAFEBASGSEKKLALOGILNNCSSLSLSPASSTL	100	
QY	137	KPMLODFPTALPARGLDHTAENTLSTYLDKSKSLCAAEVYCKEWMRYTSDGMLNKKLIERM	196	
Dd	102	DSJLRDLRDLSTLLP-----EISFRLSTFLDARSLOQAQVSKHMKELADDVDVIMHRCQEH	157	
QY	197	VRTDSLWRGLAERBGWQGYLFKN-----	219	
Dd	158	INRK-----CEKQGMGLPLLEERTLYLAANAASIQKREERTLKGVDOAHSSPYKRAKLD	211	
QY	220	-----KPPDGNAPPNNSF-----YRALYPKIIDIETIESNMRCGRHSLO	258	
Dd	212	DYPTSSNEETISSVKPPSPNSDSKFEFLPFKTRPMKEVYAEBCR-----VECNMWRHGR----	263	
QY	259	RIHR-----SETSGVYCLQDOOKIYSGSLDNTIKIIMDNTILECKRIILGHGHSVYCLQ	314	
Dd	264	-----CROVYLSGSHDGVKQCLQVRNITLASSGYDATIRLMLNATFQOVALBGHSHSGVYCLQ	320	
QY	315	YDERVIITGSSDSTVVRWADVNTGEMLTLLIHCEAVLHREFNNGMWTGSKDRSIAVWDM	374	
Dd	321	FDQCKLSISMDKTIIRIMNRYTSCISILHGHDSVCLNFEDSTLLVSGSADCVKRAMHF	380	
QY	375	ASPDITLRLRVLVGHRAVNVVDF--DDKIVASGGRITKWNSTSCFCEVRYPLINGHKRG	432	
Dd	381	SGGRITLRL-----GHTEPVNSVRIIRIRGVLVSGSDSTIKIWSLENTLCHFEFSAHIGP	436	
QY	433	IACLOYRDLRVAGSSSDNTIRLMDIEGCAQLRYEGHEEIVRCIRFENKRIRVSGAYGKI	492	
Dd	437	VQSLADLADSRFLSCGSLDGTIKQMDIEKKCKVHTLFGHIEGWEIADHLMLISAGHGV	496	
QY	493	KWMDLVALLDRAPAGTLCRLTEVHSGRVRIQFDEFQIYSSSHDPTILIMDFLNDP	550	
Dd	497	KWPMACE-----CVHTLKNBHEPYSVALGCEVAVSGSEDECKIYLMLEFNAP	543	

```

RESULT      5
S49932
MET30 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein Y19905.02; protein YIL046w
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 26-May-2000
C:Accession: S49932; S43750
R:Odell, C.; Bowman, S.
submitted to the EMBL Data Library, December 1994
A:Reference number: S49931
A:Accession: S49932
A:Molecule type: DNA
A:Residues: 1-640 <CODE>
A:Cross-references: CB:Z47047; EMBL:Z46861; NID:g603997;
R:Thomas, D.; Cherst, H.; Bailey, R.; Surdin-Kerjan, Y.
submitted to the EMBL Data Library, December 1993
A:Reference number: S43750
A:Accession: S43750

```

A:Molecule type: DNA
A:Residues: 1-60, 'I', 62-640 <THO>

A;Cross-references: EMBL:L26505; NID:g432493; PID:g432494

C;Genetics:

A: Gene: SCD:MET30

A: Cross-references: SGD.S0001308: MTPS.YT1.046W

A: Man position: 9

A;map position: 9L

C;Superfamily: unassigned WD repeat protein

F;298-329/Domain: WD repeat homology <WD1>

F;338-369/Domain: WD repeat homology <WD2>

F:374-409/Domain: WD repeat homology <WD3>

Query Match	18.0%;	Score 545;	DB 2;	length 640;
Best Local Similarly	28.9%;	Pred. No. 2.4e-34;		
Matches 167;	Conservative 85;	Mismatches 202;	Indels 124;	Gaps 17;

```

QY 34 IPEKNSLQUTNOSCARCLN--QEVVCLASTAMKENCNVAR--TKLANSTSMIYPOKJIS 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 LPEINFTKCYRANHPDIOFSTPHACTYKODLKRIQOEINAINAKIPLDQSDIHIIISKIS 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 92 ASYEKEKELCVKYEOWSESDOVEVEHLISOMCHYONGHINSYKLEMDLORDETALPAR 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 NSNDKIRKL-----ILDGIIISTSCFPLSYSSILVTHMIKIDFISILP-- 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 GLDHAENILSTDAKSTCAAEIYCKEMRYTSDGMIMKKLIERMVTBDSLMKGLAERR- 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 --QEISLKIISTLYDOSCISACNATRCRKMOKLADDDRWYHMCQHI-----DRKC 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 --GMOGLFENK-----PRDGNAPRSPFALYKPIODIETIESNMRGHRSLQ 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 PNCGMGLPLHMKRARIDONSTGSSSNADIODTOTTRPWKATYRERFEVESNMKRG----- 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 RHCRSETSK-----GVYCLQYDDPKIYSGLRDNTIKIMDNLTLECKRIKLTGHTSVCLQ 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 --HORIDFENKHMDOGLVLTLEFNRYLLFTGSDYSTIGIMDLFTGLKILRLSGHSQVXTY 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 315 YDERVITITGSSDSFVRWVDVTGEMLTLLHNCBAVYLHLRPNNGMAYTSGKDNSIAVMDI 374
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 FDDKRLITLGSDKTIKRWVNYITGECISTYRQHSQPSVLSVSYOKVIVYSGADTVKWHV 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 375 ASPDITLRLVGNHRAVNVYDDEDKYI--VSAAGSRITIKVMNTSCFCEVRLNCH-- 429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 410 ESRICYTLR---GHEWMCNVKILHRPSCFSCSDDTTIRMMDIRTNOSCLKVPGRHVQ 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 430 KRGIACIOLYRD--RLVYVSGSS----- 448
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 466 VQIKIPLIKIYENMLADNNTSDSSPODDPTMIDGAESEDTSPSNEQTVLDENIPRYTHL 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 449 -----DNTIRLMDIEGCAFLVEEGHEELVRCIRFDNKRIVSGAYDGKIYWDI.VAALDP 503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 526 LSCGLDNTIKIMDVKTGKCIFITQFGHNEGVDIADNFRILISGSHDSIKYMDLOSQ--- 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 504 RAPAGTLCRLTVLHSGVNFRLQDEFOIYSSHDPTI 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 583 -----KCMHTF---NGR--RLQRETOHOTQISIGDXY 609
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT      6
T46660
sulfur controller-2 protein [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
C:Accession: T46660
R:Kumar, A.; Paletta, J.V.
Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995.
A:Title: The sulfur controller-2, negative regulatory gene of Neurospora crassa encoded
A:Reference number: 223121, MIMD:95241499
A:Accession: T46660
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-650 <KIM>

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:36:01 ; Search time 13.01 Seconds

(without alignments)
1498.184 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034
Sequence: 1 MDPAAVLOEKALFKMNSSE.....PAAQAPPPSPRTTYISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3006	99.1	605	1	FW1A_HUMAN
2	2597	85.6	518	1	TRCB_XENLA
3	2384.5	78.6	542	1	FW1B_HUMAN
4	1635.5	53.9	701	1	YSS1_CAEEL
5	590.5	19.5	605	1	POF1_SCHPO
6	575	19.0	678	1	SCOB_EMENT
7	545	18.0	640	1	MT30_YEAST
8	531.5	17.5	650	1	SCQ2_YEUCR
9	520	17.1	579	1	SE10_CAEEL
10	519.5	17.1	267	1	YAF1_SCHPO
11	477.5	15.7	684	1	CC4_CANAL
12	455.5	15.0	1356	1	HEP1_PODAN
13	453	14.9	775	1	POPI_SCHPO
14	399	13.2	779	1	CC4_YEAST
15	396.5	13.1	703	1	POP2_SCHPO
16	392	12.9	732	1	KHMB_DICDI
17	354	11.7	409	1	LISI_HUMAN
18	354	11.7	409	1	LISI_MOUSE
19	353	11.6	409	1	LISI_BOVIN
20	341	11.2	515	1	YCM2_YEAST
21	337	11.1	1146	1	KHMA_DICDI
22	334.5	11.0	422	1	FMW2_HUMAN
23	325.5	10.7	361	1	WDS_DROME
24	324.5	10.7	362	1	WDS_HUMAN
25	318	10.5	422	1	FMW2_MOUSE
26	313.5	10.3	376	1	YKX4_CAEEL
27	312.5	10.3	714	1	YDL2_YEAST
28	307.5	10.1	742	1	PKWA_THECU
29	307.5	10.1	1693	1	Y163_SYNZ3
30	306	10.1	704	1	T2D4_DROME
31	299.5	9.9	1194	1	APAF_HUMAN
32	298.5	9.8	800	1	T2D4_HUMAN
33	292	9.6	659	1	YK16_YEAST

34	290.5	9.6	614	1	VA3A_SCHPO	Q09715 schizosacch
35	287.5	9.5	327	1	GLP1_BRANA	Q39336 brassica na
36	285.5	9.4	798	1	T2D4_YEAST	P38129 saccharomyc
37	283.5	9.3	327	1	GLP1_ARATH	Q24456 arabidopsis
38	283	9.3	473	1	PRL1_SCHPO	Q13615 schizosacch
39	283	9.3	713	1	TUPL1_YEAST	P16649 saccharomyc
40	282	9.3	318	1	GLP1_YEAST	O18640 drosophila
41	281.5	9.3	682	1	TUPL1_KLULA	P56094 kluyveromyc
42	280.5	9.2	444	1	NIDE_EMENT	Q00664 emericella
43	276.5	9.1	501	1	YH92_CAEEL	Q23256 caenorhabdi
44	273	9.0	325	1	GLP1_MEDSA	Q24076 medicago sa
45	270	8.9	451	1	PRL1_YEAST	Q12417 saccharomyc

ALIGNMENTS

RESULT	ID	FW1A_HUMAN	STANDARD:	PRT:	605 AA.
AC	Q9Y297	Q9Y213:			
DT	01-OCT-2000	(Rel. 40, Created)			
DT	01-OCT-2000	(Rel. 40, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP)				
DE	(E3RSIKAPPAB) (PIKAPPAALPHA-E3 RECEPTOR SUBUNIT).				
GN	FBXW1A OR FBW1A OR BTRCP OR BTRC.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=99075339; PubMed=985996;				
RA	Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,				
RT	Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,				
RT	"Identification of the receptor component of the I-kappaB-alpha-				
RT	ubiquitin ligase."				
RL	Nature 396:590-594(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RX	TISSUE=Lymphoid;				
RA	MEDLINE=98325370; PubMed=9660940;				
RA	Margotlin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,				
RA	Thomas D., Strebel K., Benarous R.;				
RT	"A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu				
RT	connects CD4 to the ER degradation pathway through an F-box motif."				
RL	Mol. Cell 1:565-574(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RA	Cenciarelli C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M.,				
RA	Pagano M.;				
RT	"Identification of a family of human F-box proteins."				
RL	Curr. Biol. 9:1177-1179(1999).				
RN	[4]				
RP	CHARACTERIZATION.				
RA	Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,				
RA	Harper J.W.;				
RT	"The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically				
RT	with phosphorylated destruction motifs in I-kappa-B-alpha and				
RT	beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro."				
RL	Genes Dev. 13:270-283(1999).				
CC	-1- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA				
CC	(PIKAPPAALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR				
CC	UBQUITINATION AND DEGRADATION.				
CC	-1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.				
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.				
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE				
CC	PRODUCED BY ALTERNATIVE SPLICING.				
CC	-1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.				
CC	-1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF101784; AAD08702.1; -
 DR EMBL: Y14153; CAA74572.1; -
 DR EMBL: AF129530; AAF04464.1; -
 DR MIM: 603482; -
 DR InterPro: IPR001680; -
 DR InterPro: IPR001810; -
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ubiquitin conjugation; Repeat; WD repeat; Alternative splicing.
 FT DOMAIN 190 228 F-box.
 FT REPEAT 301 338 WD 1.
 FT REPEAT 341 378 WD 2.
 FT REPEAT 381 418 WD 3.
 FT REPEAT 424 461 WD 4.
 FT REPEAT 464 503 WD 5.
 FT REPEAT 505 541 WD 6.
 FT REPEAT 553 590 WD 7.
 FT VARSPLIC 17 52 MISSING (IN ISIFORM 2).
 SQ SEQUENCE 605 AA; 68866 MW; 4C67F3B7E40ED37 CRC64;

Query Match 99.1%; Score 3006; DB 1: Length 605;
 Best Local Similarity 94.0%; Pred. No. 1.4e-215;
 Matches 369; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDAEAVLDEKALFEM-----NSSREDC 24
 DB 1 MDAEAVLDEKALFEMCSMPRLSLGSSSLADSMPSRLCLYNPGTALAFQNSSREDC 60
 QY 25 NNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLASTAKTENCVAKTRLANGTSMIV 84
 DB 61 NNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLASTAKTENCVAKTRLANGTSMIV 120
 QY 85 PKRKLSASERKEKLCVYFEQMSRSDVEFEHLISQCHQHGHINSYLKPMQRDF 144
 DB 121 PKRKLSASERKEKLCVYFEQMSRSDVEFEHLISQCHQHGHINSYLKPMQRDF 180
 QY 145 ITPALPARGDHIADENILSYLDAKSLCAAEIYCKEYRVTSQDGLMKKLIERNVRTSLMR 204
 DB 181 ITPALPARGDHIADENILSYLDAKSLCAAEIYCKEYRVTSQDGLMKKLIERNVRTSLMR 240
 QY 205 GLAERNGWGYLFKNRPPDGNAPNSFYALYKIIODIETTESNRCGRHSLSQIRHCS 264
 DB 241 GLAERNGWGYLFKNRPPDGNAPNSFYALYKIIODIETTESNRCGRHSLSQIRHCS 300
 QY 265 ETSKGYVCLQYDDOKIVSGLRONTIKIMPKNTLECKRIITGHGVSGLCYDERVITGS 324
 DB 301 ETSKGYVCLQYDDOKIVSGLRONTIKIMPKNTLECKRIITGHGVSGLCYDERVITGS 360
 QY 325 SDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAYVTSKDRSLAVWDMASPTDITLRR 384
 DB 361 SDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAYVTSKDRSLAVWDMASPTDITLRR 420
 QY 385 VLVGHRAAVNVDFDKYIVSASGDRITIVWNTSTCEVYRTLNGHKGRIACIQYRDLVY 444
 DB 421 VLVGHRAAVNVDFDKYIVSASGDRITIVWNTSTCEVYRTLNGHKGRIACIQYRDLVY 480
 QY 445 SGSSDNTIRLMDIEGACRLVLEGHEILVRCIRFQNKRIVSGAYDGKIKVMDLVALLDR 504
 DB 481 SGSSDNTIRLMDIEGACRLVLEGHEILVRCIRFQNKRIVSGAYDGKIKVMDLVALLDR 540

QY 505 APAGTLCRTLVHSGRVRLQFDFEQIVSSSHDDTLIMDFLNDPAQAQEPSPSRPT 564
 DB 541 APAGTLCRTLVHSGRVRLQFDFEQIVSSSHDDTLIMDFLNDPAQAQEPSPSRPT 600
 QY 565 TYISR 569
 DB 601 TYISR 605

RESULT 2
 TRCB_XENLA STANDARD: PRT: 518 AA.
 AC Q91854; P70037; P70038;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BETA-TRCP (BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN).
 GN FBXW1 OR BTCP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RX MEDLINE=93330289; PubMed=8393141;
 RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
 RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
 RT anaphase are rescued by Xenopus CDNs encoding N-ras or a protein
 RT with beta-transducin repeats."
 RL Mol. Cell. Biol. 13:4953-4966(1993).
 RN [2]
 RP MEDLINE=97109804; PubMed=8952061;
 RA Hudson J.W., Alarcon V.B., Elinson R.P.;
 RT "Identification of new localized RNAs in the Xenopus oocyte by
 RT differential display PCR."
 RL Dev. Genet. 19:190-198(1996).
 CC -1- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHORYLATED
 CC PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
 CC (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
 CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
 CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
 CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBD, OR
 CC TADPOLE EMBRYO.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: M98268; AAA02810.1; -
 DR EMBL: U63921; AAB49671.1; -
 DR EMBL: U63922; AAB49672.1; -
 DR InterPro: IPR001680; -
 DR InterPro: IPR001810; -
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ubiquitin conjugation; Repeat; WD repeat.
 FM DOMAIN 119 157 F-box.

```

FT REPEAT 230 258 WD 1.
FT REPEAT 270 298 WD 2.
FT REPEAT 310 338 WD 3.
FT REPEAT 353 381 WD 4.
FT REPEAT 393 421 WD 5.
FT REPEAT 433 461 WD 6.
FT REPEAT 482 510 WD 7.
FT REPEAT 516 544 WD 8.
FT REPEAT 582 610 WD 9.
FT REPEAT 622 650 WD 10.
FT REPEAT 660 688 WD 11.
FT REPEAT 718 746 WD 12.
FT REPEAT 788 816 WD 13.
FT REPEAT 858 886 WD 14.
FT REPEAT 918 946 WD 15.
FT REPEAT 978 1006 WD 16.
FT REPEAT 1038 1066 WD 17.
FT REPEAT 1098 1126 WD 18.
FT REPEAT 1158 1186 WD 19.
FT REPEAT 1218 1246 WD 20.
FT REPEAT 1278 1306 WD 21.
FT REPEAT 1338 1366 WD 22.
FT REPEAT 1398 1426 WD 23.
FT REPEAT 1458 1486 WD 24.
FT REPEAT 1518 1546 WD 25.
FT REPEAT 1578 1606 WD 26.
FT REPEAT 1638 1666 WD 27.
FT REPEAT 1698 1726 WD 28.
FT REPEAT 1758 1786 WD 29.
FT REPEAT 1818 1846 WD 30.
FT REPEAT 1878 1906 WD 31.
FT REPEAT 1938 1966 WD 32.
FT REPEAT 1998 2026 WD 33.
FT REPEAT 2058 2086 WD 34.
FT REPEAT 2118 2146 WD 35.
FT REPEAT 2178 2206 WD 36.
FT REPEAT 2238 2266 WD 37.
FT REPEAT 2298 2326 WD 38.
FT REPEAT 2358 2386 WD 39.
FT REPEAT 2418 2446 WD 40.
FT REPEAT 2478 2506 WD 41.
FT REPEAT 2538 2566 WD 42.
FT REPEAT 2598 2626 WD 43.
FT REPEAT 2658 2686 WD 44.
FT REPEAT 2718 2746 WD 45.
FT REPEAT 2778 2806 WD 46.
FT REPEAT 2838 2866 WD 47.
FT REPEAT 2898 2926 WD 48.
FT REPEAT 2958 2986 WD 49.
FT REPEAT 3018 3046 WD 50.
FT REPEAT 3078 3106 WD 51.
FT REPEAT 3138 3166 WD 52.
FT REPEAT 3198 3226 WD 53.
FT REPEAT 3258 3286 WD 54.
FT REPEAT 3318 3346 WD 55.
FT REPEAT 3378 3406 WD 56.
FT REPEAT 3438 3466 WD 57.
FT REPEAT 3498 3526 WD 58.
FT REPEAT 3558 3586 WD 59.
FT REPEAT 3618 3646 WD 60.
FT REPEAT 3678 3706 WD 61.
FT REPEAT 3738 3766 WD 62.
FT REPEAT 3798 3826 WD 63.
FT REPEAT 3858 3886 WD 64.
FT REPEAT 3918 3946 WD 65.
FT REPEAT 3978 4006 WD 66.
FT REPEAT 4038 4066 WD 67.
FT REPEAT 4098 4126 WD 68.
FT REPEAT 4158 4186 WD 69.
FT REPEAT 4218 4246 WD 70.
FT REPEAT 4278 4306 WD 71.
FT REPEAT 4338 4366 WD 72.
FT REPEAT 4398 4426 WD 73.
FT REPEAT 4458 4486 WD 74.
FT REPEAT 4518 4546 WD 75.
FT REPEAT 4578 4606 WD 76.
FT REPEAT 4638 4666 WD 77.
FT REPEAT 4698 4726 WD 78.
FT REPEAT 4758 4786 WD 79.
FT REPEAT 4818 4846 WD 80.
FT REPEAT 4878 4906 WD 81.
FT REPEAT 4938 4966 WD 82.
FT REPEAT 4998 5026 WD 83.
FT REPEAT 5058 5086 WD 84.
FT REPEAT 5118 5146 WD 85.
FT REPEAT 5178 5206 WD 86.
FT REPEAT 5238 5266 WD 87.
FT REPEAT 5298 5326 WD 88.
FT REPEAT 5358 5386 WD 89.
FT REPEAT 5418 5446 WD 90.
FT REPEAT 5478 5506 WD 91.
FT REPEAT 5538 5566 WD 92.
FT REPEAT 5598 5626 WD 93.
FT REPEAT 5658 5686 WD 94.
FT REPEAT 5718 5746 WD 95.
FT REPEAT 5778 5806 WD 96.
FT REPEAT 5838 5866 WD 97.
FT REPEAT 5898 5926 WD 98.
FT REPEAT 5958 5986 WD 99.
FT REPEAT 6018 6046 WD 100.
FT REPEAT 6078 6106 WD 101.
FT REPEAT 6138 6166 WD 102.
FT REPEAT 6198 6226 WD 103.
FT REPEAT 6258 6286 WD 104.
FT REPEAT 6318 6346 WD 105.
FT REPEAT 6378 6406 WD 106.
FT REPEAT 6438 6466 WD 107.
FT REPEAT 6498 6526 WD 108.
FT REPEAT 6558 6586 WD 109.
FT REPEAT 6618 6646 WD 110.
FT REPEAT 6678 6706 WD 111.
FT REPEAT 6738 6766 WD 112.
FT REPEAT 6798 6826 WD 113.
FT REPEAT 6858 6886 WD 114.
FT REPEAT 6918 6946 WD 115.
FT REPEAT 6978 7006 WD 116.
FT REPEAT 7038 7066 WD 117.
FT REPEAT 7098 7126 WD 118.
FT REPEAT 7158 7186 WD 119.
FT REPEAT 7218 7246 WD 120.
FT REPEAT 7278 7306 WD 121.
FT REPEAT 7338 7366 WD 122.
FT REPEAT 7398 7426 WD 123.
FT REPEAT 7458 7486 WD 124.
FT REPEAT 7518 7546 WD 125.
FT REPEAT 7578 7606 WD 126.
FT REPEAT 7638 7666 WD 127.
FT REPEAT 7698 7726 WD 128.
FT REPEAT 7758 7786 WD 129.
FT REPEAT 7818 7846 WD 130.
FT REPEAT 7878 7906 WD 131.
FT REPEAT 7938 7966 WD 132.
FT REPEAT 7998 8026 WD 133.
FT REPEAT 8058 8086 WD 134.
FT REPEAT 8118 8146 WD 135.
FT REPEAT 8178 8206 WD 136.
FT REPEAT 8238 8266 WD 137.
FT REPEAT 8298 8326 WD 138.
FT REPEAT 8358 8386 WD 139.
FT REPEAT 8418 8446 WD 140.
FT REPEAT 8478 8506 WD 141.
FT REPEAT 8538 8566 WD 142.
FT REPEAT 8598 8626 WD 143.
FT REPEAT 8658 8686 WD 144.
FT REPEAT 8718 8746 WD 145.
FT REPEAT 8778 8806 WD 146.
FT REPEAT 8838 8866 WD 147.
FT REPEAT 8898 8926 WD 148.
FT REPEAT 8958 8986 WD 149.
FT REPEAT 9018 9046 WD 150.
FT REPEAT 9078 9106 WD 151.
FT REPEAT 9138 9166 WD 152.
FT REPEAT 9198 9226 WD 153.
FT REPEAT 9258 9286 WD 154.
FT REPEAT 9318 9346 WD 155.
FT REPEAT 9378 9406 WD 156.
FT REPEAT 9438 9466 WD 157.
FT REPEAT 9498 9526 WD 158.
FT REPEAT 9558 9586 WD 159.
FT REPEAT 9618 9646 WD 160.
FT REPEAT 9678 9706 WD 161.
FT REPEAT 9738 9766 WD 162.
FT REPEAT 9798 9826 WD 163.
FT REPEAT 9858 9886 WD 164.
FT REPEAT 9918 9946 WD 165.
FT REPEAT 9978 10006 WD 166.
FT REPEAT 10038 10066 WD 167.
FT REPEAT 10098 10126 WD 168.
FT REPEAT 10158 10186 WD 169.
FT REPEAT 10218 10246 WD 170.
FT REPEAT 10278 10306 WD 171.
FT REPEAT 10338 10366 WD 172.
FT REPEAT 10398 10426 WD 173.
FT REPEAT 10458 10486 WD 174.
FT REPEAT 10518 10546 WD 175.
FT REPEAT 10578 10606 WD 176.
FT REPEAT 10638 10666 WD 177.
FT REPEAT 10698 10726 WD 178.
FT REPEAT 10758 10786 WD 179.
FT REPEAT 10818 10846 WD 180.
FT REPEAT 10878 10906 WD 181.
FT REPEAT 10938 10966 WD 182.
FT REPEAT 10998 11026 WD 183.
FT REPEAT 11058 11086 WD 184.
FT REPEAT 11118 11146 WD 185.
FT REPEAT 11178 11206 WD 186.
FT REPEAT 11238 11266 WD 187.
FT REPEAT 11298 11326 WD 188.
FT REPEAT 11358 11386 WD 189.
FT REPEAT 11418 11446 WD 190.
FT REPEAT 11478 11506 WD 191.
FT REPEAT 11538 11566 WD 192.
FT REPEAT 11598 11626 WD 193.
FT REPEAT 11658 11686 WD 194.
FT REPEAT 11718 11746 WD 195.
FT REPEAT 11778 11806 WD 196.
FT REPEAT 11838 11866 WD 197.
FT REPEAT 11898 11926 WD 198.
FT REPEAT 11958 11986 WD 199.
FT REPEAT 12018 12046 WD 200.
FT REPEAT 12078 12106 WD 201.
FT REPEAT 12138 12166 WD 202.
FT REPEAT 12198 12226 WD 203.
FT REPEAT 12258 12286 WD 204.
FT REPEAT 12318 12346 WD 205.
FT REPEAT 12378 12406 WD 206.
FT REPEAT 12438 12466 WD 207.
FT REPEAT 12498 12526 WD 208.
FT REPEAT 12558 12586 WD 209.
FT REPEAT 12618 12646 WD 210.
FT REPEAT 12678 12706 WD 211.
FT REPEAT 12738 12766 WD 212.
FT REPEAT 12798 12826 WD 213.
FT REPEAT 12858 12886 WD 214.
FT REPEAT 12918 12946 WD 215.
FT REPEAT 12978 12986 WD 216.
FT REPEAT 13038 13066 WD 217.
FT REPEAT 13098 13126 WD 218.
FT REPEAT 13158 13186 WD 219.
FT REPEAT 13218 13246 WD 220.
FT REPEAT 13278 13306 WD 221.
FT REPEAT 13338 13366 WD 222.
FT REPEAT 13398 13426 WD 223.
FT REPEAT 13458 13486 WD 224.
FT REPEAT 13518 13546 WD 225.
FT REPEAT 13578 13606 WD 226.
FT REPEAT 13638 13666 WD 227.
FT REPEAT 13698 13726 WD 228.
FT REPEAT 13758 13786 WD 229.
FT REPEAT 13818 13846 WD 230.
FT REPEAT 13878 13906 WD 231.
FT REPEAT 13938 13966 WD 232.
FT REPEAT 13998 14026 WD 233.
FT REPEAT 14058 14086 WD 234.
FT REPEAT 14118 14146 WD 235.
FT REPEAT 14178 14206 WD 236.
FT REPEAT 14238 14266 WD 237.
FT REPEAT 14298 14326 WD 238.
FT REPEAT 14358 14386 WD 239.
FT REPEAT 14418 14446 WD 240.
FT REPEAT 14478 14506 WD 241.
FT REPEAT 14538 14566 WD 242.
FT REPEAT 14598 14626 WD 243.
FT REPEAT 14658 14686 WD 244.
FT REPEAT 14718 14746 WD 245.
FT REPEAT 14778 14806 WD 246.
FT REPEAT 14838 14866 WD 247.
FT REPEAT 14898 14926 WD 248.
FT REPEAT 14958 14986 WD 249.
FT REPEAT 15018 15046 WD 250.
FT REPEAT 15078 15106 WD 251.
FT REPEAT 15138 15166 WD 252.
FT REPEAT 15198 15226 WD 253.
FT REPEAT 15258 15286 WD 254.
FT REPEAT 15318 15346 WD 255.
FT REPEAT 15378 15406 WD 256.
FT REPEAT 15438 15466 WD 257.
FT REPEAT 15498 15526 WD 258.
FT REPEAT 15558 15586 WD 259.
FT REPEAT 15618 15646 WD 260.
FT REPEAT 15678 15706 WD 261.
FT REPEAT 15738 15766 WD 262.
FT REPEAT 15798 15826 WD 263.
FT REPEAT 15858 15886 WD 264.
FT REPEAT 15918 15946 WD 265.
FT REPEAT 15978 16006 WD 266.
FT REPEAT 16038 16066 WD 267.
FT REPEAT 16098 16126 WD 268.
FT REPEAT 16158 16186 WD 269.
FT REPEAT 16218 16246 WD 270.
FT REPEAT 16278 16306 WD 271.
FT REPEAT 16338 16366 WD 272.
FT REPEAT 16398 16426 WD 273.
FT REPEAT 16458 16486 WD 274.
FT REPEAT 16518 16546 WD 275.
FT REPEAT 16578 16606 WD 276.
FT REPEAT 16638 16666 WD 277.
FT REPEAT 16698 16726 WD 278.
FT REPEAT 16758 16786 WD 279.
FT REPEAT 16818 16846 WD 280.
FT REPEAT 16878 16906 WD 281.
FT REPEAT 16938 16966 WD 282.
FT REPEAT 16998 17026 WD 283.
FT REPEAT 17058 17086 WD 284.
FT REPEAT 17118 17146 WD 285.
FT REPEAT 17178 17206 WD 286.
FT REPEAT 17238 17266 WD 287.
FT REPEAT 17298 17326 WD 288.
FT REPEAT 17358 17386 WD 289.
FT REPEAT 17418 17446 WD 290.
FT REPEAT 17478 17506 WD 291.
FT REPEAT 17538 17566 WD 292.
FT REPEAT 17598 17626 WD 293.
FT REPEAT 17658 17686 WD 294.
FT REPEAT 17718 17746 WD 295.
FT REPEAT 17778 17806 WD 296.
FT REPEAT 17838 17866 WD 297.
FT REPEAT 17898 17926 WD 298.
FT REPEAT 17958 17986 WD 299.
FT REPEAT 18018 18046 WD 300.
FT REPEAT 18078 18106 WD 301.
FT REPEAT 18138 18166 WD 302.
FT REPEAT 18198 18226 WD 303.
FT REPEAT 18258 18286 WD 304.
FT REPEAT 18318 18346 WD 305.
FT REPEAT 18378 18406 WD 306.
FT REPEAT 18438 18466 WD 307.
FT REPEAT 18498 18526 WD 308.
FT REPEAT 18558 18586 WD 309.
FT REPEAT 18618 18646 WD 310.
FT REPEAT 18678 18706 WD 311.
FT REPEAT 18738 18766 WD 312.
FT REPEAT 18798 18826 WD 313.
FT REPEAT 18858 18886 WD 314.
FT REPEAT 18918 18946 WD 315.
FT REPEAT 18978 19006 WD 316.
FT REPEAT 19038 19066 WD 317.
FT REPEAT 19098 19126 WD 318.
FT REPEAT 19158 19186 WD 319.
FT REPEAT 19218 19246 WD 320.
FT REPEAT 19278 19306 WD 321.
FT REPEAT 19338 19366 WD 322.
FT REPEAT 19398 19426 WD 323.
FT REPEAT 19458 19486 WD 324.
FT REPEAT 19518 19546 WD 325.
FT REPEAT 19578 19606 WD 326.
FT REPEAT 19638 19666 WD 327.
FT REPEAT 19698 19726 WD 328.
FT REPEAT 19758 19786 WD 329.
FT REPEAT 19818 19846 WD 330.
FT REPEAT 19878 19906 WD 331.
FT REPEAT 19938 19966 WD 332.
FT REPEAT 19998 20026 WD 333.
FT REPEAT 20058 20086 WD 334.
FT REPEAT 20118 20146 WD 335.
FT REPEAT 20178 20206 WD 336.
FT REPEAT 20238 20266 WD 337.
FT REPEAT 20298 20326 WD 338.
FT REPEAT 20358 20386 WD 339.
FT REPEAT 20418 20446 WD 340.
FT REPEAT 20478 20506 WD 341.
FT REPEAT 20538 20566 WD 342.
FT REPEAT 20598 20626 WD 343.
FT REPEAT 20658 20686 WD 344.
FT REPEAT 20718 20746 WD 345.
FT REPEAT 20778 20806 WD 346.
FT REPEAT 20838 20866 WD 347.
FT REPEAT 20898 20926 WD 348.
FT REPEAT 20958 20986 WD 349.
FT REPEAT 21018 21046 WD 350.
FT REPEAT 21078 21106 WD 351.
FT REPEAT 21138 21166 WD 352.
FT REPEAT 21198 21226 WD 353.
FT REPEAT 21258 21286 WD 354.
FT REPEAT 21318 21346 WD 355.
FT REPEAT 21378 21406 WD 356.
FT REPEAT 21438 21466 WD 357.
FT REPEAT 21498 21526 WD 358.
FT REPEAT 21558 21586 WD 359.
FT REPEAT 21618 21646 WD 360.
FT REPEAT 21678 21706 WD 361.
FT REPEAT 21738 21766 WD 362.
FT REPEAT 21798 21826 WD 363.
FT REPEAT 21858 21886 WD 364.
FT REPEAT 21918 21946 WD 365.
FT REPEAT 21978 22006 WD 366.
FT REPEAT 22038 22066 WD 367.
FT REPEAT 22098 22126 WD 368.
FT REPEAT 22158 22186 WD 369.
FT REPEAT 22218 22246 WD 370.
FT REPEAT 22278 22306 WD 371.
FT REPEAT 22338 22366 WD 372.
FT REPEAT 22398 22426 WD 373.
FT REPEAT 22458 22486 WD 374.
FT REPEAT 22518 22546 WD 375.
FT REPEAT 22578 22606 WD 376.
FT REPEAT 22638 22666 WD 377.
FT REPEAT 22698 22726 WD 378.
FT REPEAT 22758 22786 WD 379.
FT REPEAT 22818 22846 WD 380.
FT REPEAT 22878 22906 WD 381.
FT REPEAT 22938 22966 WD 382.
FT REPEAT 22998 23026 WD 383.
FT REPEAT 23058 23086 WD 384.
FT REPEAT 23118 23146 WD 385.
FT REPEAT 23178 23206 WD 386.
FT REPEAT 23238 23266 WD 387.
FT REPEAT 23298 23326 WD 388.
FT REPEAT 23358 23386 WD 389.
FT REPEAT 23418 23446 WD 390.
FT REPEAT 23478 23506 WD 391.
FT REPEAT 23538 23566 WD 392.
FT REPEAT 23598 23626 WD 393.
FT REPEAT 23658 23686 WD 394.
FT REPEAT 23718 23746 WD 395.
FT REPEAT 23778 23806 WD 396.
FT REPEAT 23838 23866 WD 397.
FT REPEAT 23898 23926 WD 398.
FT REPEAT 23958 23986 WD 399.
FT REPEAT 24018 24046 WD 400.
FT REPEAT 24078 24106 WD 401.
FT REPEAT 24138 24166 WD 402.
FT REPEAT 24198 24226 WD 403.
FT REPEAT 24258 24286 WD 404.
FT REPEAT 24318 24346 WD 405.
FT REPEAT 24378 24406 WD 406.
FT REPEAT 24438 24466 WD 407.
FT REPEAT 24498 24526 WD 408.
FT REPEAT 24558 24586 WD 409.
FT REPEAT 24618 24646 WD 410.
FT REPEAT 24678 24706 WD 411.
FT REPEAT 24738 24766 WD 412.
FT REPEAT 24798 24826 WD 413.
FT REPEAT 24858 24886 WD 414.
FT REPEAT 24918 24946 WD 415.
FT REPEAT 24978 25006 WD 416.
FT REPEAT 25038 25066 WD 417.
FT REPEAT 25098 25126 WD 418.
FT REPEAT 25158 25186 WD 419.
FT REPEAT 25218 25246 WD 420.
FT REPEAT 25278 25306 WD 421.
FT REPEAT 25338 25366 WD 422.
FT REPEAT 25398 25426 WD 423.
FT REPEAT 25458 25486 WD 424.
FT REPEAT 25518 25546 WD 425.
FT REPEAT 25578 25606 WD 426.
FT REPEAT 25638 25666 WD 427.
FT REPEAT 25698 25726 WD 428.
FT REPEAT 25758 25786 WD 429.
FT REPEAT 25818 25846 WD 430.
FT REPEAT 25878 25906 WD 431.
FT REPEAT 25938 25966 WD 432.
FT REPEAT 25998 26026 WD 433.
FT REPEAT 26058 26086 WD 434.
FT REPEAT 26118 26146 WD 435.
FT REPEAT 26178 26206 WD 436.
FT REPEAT 26238 26266 WD 437.
FT REPEAT 26298 26326 WD 438.
FT REPEAT 26358 26386 WD 439.
FT REPEAT 26418 26446 WD 440.
FT REPEAT 26478 26506 WD 441.
FT REPEAT 26538 26566 WD 442.
FT REPEAT 26598 26626 WD 443.
FT REPEAT 26658 26686 WD 444.
FT REPEAT 26718 26746 WD 445.
FT REPEAT 26778 26806 WD 446.
FT REPEAT 26838 26866 WD 447.
FT REPEAT 26898 26926 WD 448.
FT REPEAT 26958 26986 WD 449.
FT REPEAT 27018 27046 WD 450.
FT REPEAT 27078 27106 WD 451.
FT REPEAT 27138 27166 WD 452.
FT REPEAT 27198 27226 WD 453.
FT REPEAT 27258 27286 WD 454.
FT REPEAT 27318 27346 WD 455.
FT REPEAT 27378 27406 WD 456.
FT REPEAT 27438 27466 WD 457.
FT REPEAT 27498 27526 WD 458.
FT REPEAT 27558 27586 WD 459.
FT REPEAT 27618 27646 WD 460.
FT REPEAT 27678 27706 WD 461.
FT REPEAT 27738 27766 WD 462.
FT REPEAT 27798 27826 WD 463.
FT REPEAT 27858 27886 WD 464.
FT REPEAT 27918 27946 WD 465.
FT REPEAT 27978 27986 WD 466.
FT REPEAT 28038 28066 WD 467.
FT REPEAT 28098 28126 WD 468.
FT REPEAT 28158 28186 WD 469.
FT REPEAT 28218 28246 WD 470.
FT REPEAT 28278 28306 WD 471.
FT REPEAT 28338 28366 WD 472.
FT REPEAT 28398 28426 WD 473.
FT REPEAT 28458 28486 WD 474.
FT REPEAT 28518 28546 WD 475.
FT REPEAT 28578 28606 WD 476.
FT REPEAT 28638 28666 WD 477.
FT REPEAT 28698 28726 WD 478.
FT REPEAT 28758 28786 WD 479.
FT REPEAT 28818 28846 WD 480.
FT REPEAT 28878 28906 WD 481.
FT REPEAT 28938 28966 WD 482.
FT REPEAT 28998 29026 WD 483.
FT REPEAT 29058 29086 WD 484.
FT REPEAT 29118 29146 WD 485.
FT REPEAT 29178 29206 WD 486.
FT REPEAT 29238 29266 WD 487.
FT REPEAT 29298 29326 WD 488.
FT REPEAT 29358 29386 WD 489.
FT REPEAT 29418 29446 WD 490.
FT REPEAT 29478 29506 WD 491.
FT REPEAT 29538 29566 WD 492.
FT REPEAT 29598 29626 WD 493.
FT REPEAT 29658 29686 WD 494.
FT REPEAT 29718 29746 WD 495.
FT REPEAT 29778 29806 WD 496.
FT REPEAT 29838 29866 WD 497.
FT REPEAT 29898 29926 WD 498.
FT REPEAT 29958 29986 WD 499.
FT REPEAT 30018 30046 WD 500.
FT REPEAT 30078 30106 WD 501.
FT REPEAT 30138 30166 WD 502.
FT REPEAT 30198 30226 WD 503.
FT REPEAT 30258 30286 WD 504.
FT REPEAT 30318 30346 WD 505.
FT REPEAT 30378 30406 WD 506.
FT REPEAT 30438 30466 WD 507.
FT REPEAT 30498 30526 WD 508.
FT REPEAT 30558 30586 WD 509.
FT REPEAT 30618 30646 WD 510.
FT REPEAT 30678 30706 WD 511.
FT REPEAT 30738 30766 WD 512.
FT REPEAT 30798 30826 WD 513.
FT REPEAT 30858 30886 WD 514.
FT REPEAT 30918 30946 WD 515.
FT REPEAT 30978 31006 WD 516.
FT REPEAT 31038 31066 WD 517.
FT REPEAT 31098 31126 WD 518.
FT REPEAT 31158 31186 WD 519.
FT REPEAT 31218 31246 WD 520.
FT REPEAT 31278 31306 WD 521.
FT REPEAT 31338 31366 WD 522.
FT REPEAT 31398 31426 WD 523.
FT REPEAT 31458 31486 WD 524.
FT REPEAT 31518 31546 WD 525.
FT REPEAT 31578 31606 WD 526.
FT REPEAT 31638 31666 WD 527.
FT REPEAT 31698 31726 WD 528.
FT REPEAT 31758 31786 WD 529.
FT REPEAT 31818 31846 WD 530.
FT REPEAT 31878 31906 WD 531.
FT REPEAT 31938 31966 WD 532.
FT REPEAT 31998 32026 WD 533.
FT REPEAT 32058 32086 WD 534.
FT REPEAT 32118 32146 WD 535.
FT REPEAT 32178 32206 WD 536.
FT REPEAT 32238 32266 WD 537.
FT REPEAT 32298 32326 WD 538.
FT REPEAT 32358 32386 WD 539.
FT REPEAT 32418 32446 WD 540.
FT REPEAT 32478 32506 WD 541.
FT REPEAT 32538 32566 WD 542.
FT REPEAT 32598 32626 WD 543.
FT REPEAT 32658 32686 WD 544.
FT REPEAT 32718 32746 WD 545.
FT REPEAT 32778 32806 WD 546.
FT REPEAT 32838 32866 WD 547.
FT REPEAT 32898 32926 WD 548.
FT REPEAT 32958 32986 WD 549.
FT REPEAT 33018 33046 WD 550.
FT REPEAT 33078 33106 WD 551.
FT REPEAT 33138 33166 WD 552.
FT REPEAT 33198 33226 WD 553.
FT REPEAT 33258 33286 WD 554.
FT REPEAT 33318 33346 WD 555.
FT REPEAT 33378 33406 WD 556.
FT REPEAT 33438 33466 WD 557.
FT REPEAT 33498 33526 WD 558.
FT REPEAT 33558 33586 WD 559.
FT REPEAT 33618 33646 WD 560.
FT REPEAT 33678 33706 WD 561.
FT REPEAT 33738 33766 WD 562.
FT REPEAT 33798 33826 WD 563.
FT REPEAT 33858 33886 WD 564.
FT REPEAT 33918 33946 WD 565.
FT REPEAT 33978 33986 WD 566.
FT REPEAT 34038 34066 WD 567.
FT REPEAT 34098 34126 WD 568.
FT REPEAT 34158 34186 WD 569.
FT REPEAT 34218 34246 WD 570.
FT REPEAT 34278 34306 WD 571.
FT REPEAT 34338 34366 WD 572.
FT REPEAT 34398 34426 WD 573.
FT REPEAT 34458 34486 WD 574.
FT REPEAT 34518 34546 WD 575.
FT REPEAT 34578 34606 WD 576.
FT REPEAT 34638 34666 WD 577.

```

```

Db 37 SCLQSNPVSRYCL---OISNGTSSVIVSRKRPBGRNYQKEKDLCKYFDOMSESQYEFVE 93
QY 119 HLTSGMCHVOHGHINSYKLPMLORDFTLTPARGLDHTAENITLSTYDAKSLCAELVCKE 178
Db 94 HLTSGMCHVOHGHINSYKLPMLORDFTLTPARGLDHTAENITLSTYDAKSLCAELVCKE 153
QY 179 WYRVTSDGMLMKLLIERMVRTDLSLARGLAERRGMGOYLEFKNKPDPGNAPNSFYRALYRK 238
Db 154 WQVIVSBGMMLKLLIERMVRTDLSLARGLAERRGMGOYLEFKNKPDPGNAPNSFYRALYRK 211
QY 239 IIDDITETIESNMRCGRHSLORHCRSETSKGYCCLOYDDOKIVSGLRDNTIKTWKNTLE 298
Db 212 IIDDITETIESNMRCGRHSLORHCRSETSKGYCCLOYDDOKIVSGLRDNTIKTWKNTLE 271
QY 299 CKRLHGHGTSVCLQYDERVITITGSSDSSTVRYWMDVNTGEMTLTLHHCEAVLHLFPNNG 358
Db 272 CLVVLGHGTSVCLQYDERVITITGSSDSSTVRYWMDVNTGEMTLTLHHCEAVLHLFPNNG 331
QY 359 MMYTCSKDRSIAMVMDASPTDITLRRVLVGHRAAVNVDFDCKYIVASGSDRTIKWNTS 418
Db 332 LMYTCSKDRSIAMVMDASPTDITLRRVLVGHRAAVNVDFDCKYIVASGSDRTIKWNTS 391
QY 419 TGEFVRLTNGHKGACIACLOYRDLRVVSGSSDNTIRLMDIECGACLVLESGHELVRCIRP 478
Db 392 TGEFVRLTNGHKGACIACLOYRDLRVVSGSSDNTIRLMDIECGACLVLESGHELVRCIRP 451
QY 479 DNRIVSGAIDGKIKWMDLVAAIDPRAPAGTLCRLTLVHSGVRFLODEPOIVSSSHD 538
Db 452 DNRIVSGAIDGKIKWMDLVAAIDPRAPAGTLCRLTLVHSGVRFLODEPOIVSSSHD 511
QY 539 DTLIMDFLNDPAAOAEPPRSRTTYTISR 569
Db 512 DTLIMDFLNDPAAOAEPPRSRTTYTISR 542

RESULT 4
YSSL_CAEEL STANDARD: PRT; 701 AA.
AC 009990;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOPHOSOME II.
GN K1082.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Miller N.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: STRONG, TO X.LAEVIS BETA-TRCP.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U28730, AAA68258.1;
CC WormRep: K10B2.1; CE02008.
CC InterPro: IPR001680;
CC Pfam: PF00400; WD40; 7.
CC PRINTS: PR00320; GPROTEINBRPT.
CC PROSITE: PS50181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS_1; 5.
CC PROSITE: PS50082; WD_REPEATS_2; 7.

```

```

DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
FT REPEAT 256 284 WD 1.
FT REPEAT 296 324 WD 2.
FT REPEAT 336 364 WD 3.
FT REPEAT 379 407 WD 4.
FT REPEAT 419 447 WD 5.
FT REPEAT 459 487 WD 6.
FT REPEAT 508 536 WD 7.
FT DOMAIN 606 615 POLY-ALA.
SQ SEQUENCE 701 AA; 80320 MW; 69FA0B00F83270E3 CRC64;

Query Match
Best Local Similarity 53.9%; Score 1635.5; DB 1; Length 701;
Matches 328; Conservative 69; Mismatches 116; Indels 55; Gaps 8;

QY 31 RKIPEKNSLRQTYNSCARCLINDETVCIASTAMKTENCVAKTKLAN-----GTS 80
Db 2 RFRREGKRALKQGRARDOGSIOALTVCVST----IERCF--TAVSNIFLFTSFVSF 55
QY 81 SMIVRKOR-----KLSAYEKELCVKYFEQMSHSDQVEFEVHLISMCHYOHGHSY 135
Db 56 SFLPFSRTQJFLYSKRSFSFSEVL-----KWSHEQJLDMDKIVHRLSHYOLGKYDNF 110
QY 136 LKPMLOPFTLTPARGLDHTAENITLSTYDAKSLCAELVCKEYRVTSQDMLMKLLIER 195
Db 111 ITPMLQRFISLPLA-----HLVELILFNVNSDSLSCSEVSTSRCALARQHKLLIEK 166
QY 196 MVRTDSLARGLAERRGMGOYL-----FKNKPDPGNAPNSFYRAL 235
Db 167 NVRSDSLWMGJSEKQMPKFLNISRDMSVRRICEKFNVDVNIKRDLQJLIMHVFYSKL 226
QY 236 YPKIIDDITETIESNMRCGRHSLORHCRSETSKGYCCLOYDDOKIVSGLRDNTIKIMDKN 295
Db 227 YPKIIRIDHNIDNMKRNKNTKTRINCOSSENSKGYCCLOYDDOKIVSGLRDNTIKIMDK 286
QY 296 TLECKRIITGHGTSVCLQYDERVITITGSSDSSTVRYWMDVNTGEMTLTLHHCEAVLHLRF 355
Db 287 DYSCGRILISGHGTSVCLQYDNRVITSSSDATYAVWVDEGECKTLIHCEAVLHLRF 346
QY 356 NNGMAYTCSKDRSIAMVMDASPTDITLRRVLVGHRAAVNVDFDCKYIVASGSDRTIKW 415
Db 347 ANGIIVTCSKDRSIAMVMDVSPDITIRRVLVGHRAAVNVDFDCKYIVASGSDRTIKW 406
QY 416 NNSGTEFVRLTNGHKGACIACLOYRDLRVVSGSSDNTIRLMDIECGACLVLESGHELVRC 475
Db 407 SMDTLEFVRLTNGHKGACIACLOYRDLRVVSGSSDNTIRLMDIHSGVCLRVLESGHELVRC 466
QY 476 IREFDKRIIVSGAIDGKIKWMDLVAAIDPRAPAGTLCRLTLVHSGVRFLODEPOIVSS 535
Db 467 IREFDKRIIVSGAIDGKIKWMDLVAAIDPRALSEICLSLVGHGRVRLQFDEPOIVSS 526
QY 536 SHDDTITLIMDFLNDPAAOAEPPRSRT 563
Db 527 SHDDTITLIMDFLNDP-----PSGLPSST 549

RESULT 5
POF1_SCHPO STANDARD: PRT; 605 AA.
AC P87053;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE F-BOX/WD-REPEAT PROTEIN POF1.
GN POF1 OR SPAC57A10.05.C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Katayama S., Kitamura K., Toda T.;
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Baddock K., Churcher C.M., Wood V., Barrell B.G., Randalream M.A.;
RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AB032410; BAA84528.1; -
CC EMBL: Z94864; CAB08168.1; -
CC InterPro: IPR001680; -
CC InterPro: IPR001810; -
CC Pfam: PF00646; F-box.1.
CC Pfam: PF00400; WD40; 7.
CC PRINTS: PR00320; GPROTEINRPT.
CC PROSITE: PS50181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS_1; 2.
CC PROSITE: PS50082; WD_REPEATS_2; 7.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC Repeat: WD repeat.
CC KW DOMAIN 107 153 F-BOX.
CC FT REPEAT 271 299 WD 1.
CC FT REPEAT 311 339 WD 2.
CC FT REPEAT 350 379 WD 3.
CC FT REPEAT 390 420 WD 4.
CC FT REPEAT 432 460 WD 5.
CC FT REPEAT 472 500 WD 6.
CC FT REPEAT 510 538 WD 7.
CC FT REPEAT 510 538 WD 7.
CC SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 19.5%; Score 590.5; DB 1; Length 605;
Best Local Similarity 29.0%; Pred. No. 1.9e-36;
Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

OY 80 SSATVKKOKRLASVKEKELCVYKFEQNSE---SDQVEFEHLISOMCHYGHINSTL 136
DB 46 SSM---NHMELSGLSKSRQREAVAAAFSEASCSERKLALOGILNCSLSLTFASSTL 101
OY 137 KPLALORDFTALPARGLDHAENILSYLDKAKSLCAELVCKEYRYTSDGMLKRLIERH 196
DB 102 DSLVRLDPSLSLPLV---EISFRILSFLDKRSKQAAOYAKHKEILDADDVYIHRMCEQH 157
OY 197 VRTDSLMRGLAERGRGOYLFXN-----
DB 158 INNK-----CEKCGMGLPLENNTLYAAKASTOKRYERLTGRGVQAHSSPVYKAKLD 211
OY 220 -----KPPDGNAPNSP-----YRALYPKTIIDTETIESNMRCGRHSIQ 258
DB 212 DYEPTSSNETISSVKKPPSPNSDSKFLPFRTRPKWEYERCR---VECNMRHGR---- 263
OY 259 RHGR-----SETSKGYVCLQYDQKIVSGRNDTIKIMKNTLECKRIILNTHGSGVLCIQ 314
DB 264 ---CROVVISGSHDGYMCQQLVNRILASGSYDTIRLMNLATQOVALLEGHSYGTCLQ 320
OY 315 YDERVITIGSSDSTVAVDNTGEMNLTIHRCVAVLHLFNNGMVYTSKDRSIAVMDM 374
DB 321 FDCKRLISGMDKTIIRIMWRTSECSILHGHNDVSLCLTFDS7LLVSGSADCTVLMHMF 380
OY 375 ASPTDITLRVYLGHRAAVVDF---DDKTIYASGDRITKVNNTSCFVRLNLGHKRG 432

```

```

DB 381 SGGRITLR-----GHTGPVNSVRIIRDRGLVLSGSDSTIKTMSLETNCLHFFSAHICP 436
OY 433 IACLOYRDLVYSGSDNTIRLMDIEGACALRYLECHEEYVRCIRPDNRKIVSGAYDKI 492
DB 437 YOSLALADRLPSCSLDGTITKQMDIEKKCVHTLFFHIGVWEIADHLRLISGADHGVY 496
OY 493 KWDVLAALDPPRAPACTLCLRTLVHSGVRFRLQDFEFQIVSSSHDITLIDFLNDE 550
DB 497 KWEEACE-----CVHTLKNHSEPTVSVALDCVSGSEDEKITYLMLFNAP 543

RESULT 6
SCOB_EMENT STANDARD; PRT; 678 AA.
ID SCOB_EMENT 000659;
DR 15-JUL-1998 (rel. 36, Created)
DR 15-JUL-1998 (rel. 36, Last sequence update)
DR 01-OCT-2000 (rel. 40, Last annotation update)
DE SULFUR METABOLITE REPRESSION CONTROL PROTEIN.
GN SCOB OR MAPB1.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA Natorff R.;
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
CC REPRESSION.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MET30/SCOB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U21220; AAC15905.1; -
CC InterPro: IPR001680; -
CC InterPro: IPR001810; -
CC Pfam: PF00646; F-box.1.
CC Pfam: PF00400; WD40; 6.
CC PRINTS: PR00320; GPROTEINRPT.
CC PROSITE: PS50181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS_1; 4.
CC PROSITE: PS50082; WD_REPEATS_2; 7.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC Transcription regulation; Repeat; WD repeat.
CC KW DOMAIN 178 224 F-BOX.
CC FT REPEAT 347 375 WD 1.
CC FT REPEAT 387 415 WD 2.
CC FT REPEAT 427 455 WD 3.
CC FT REPEAT 466 496 WD 4.
CC FT REPEAT 508 543 WD 5.
CC FT REPEAT 553 595 WD 6.
CC FT REPEAT 607 635 WD 7.
CC FT REPEAT 647 675 WD 8.
CC SEQUENCE 678 AA; 76070 MW; D840D452E37BAC53 CRC64;

Query Match 19.0%; Score 575; DB 1; Length 678;
Best Local Similarity 26.9%; Pred. No. 3.1e-35;
Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;

OY 117 VEHLISOMCHYGHINSTYKPMALORDFTALPARGLDHAENILSYLDKAKSLCAELVLC 176
DB 153 LOGIMACCPOLSYISATVAVRDLRIDFTALP-----PEIAFKILCYLDPTSLCKASOVS 208

```

```

OY 177 KEMVRYTSDGMLKKLIERAVRTDSLRLGLAERGGQYLFKNK----- 220
D 209 RGRRLADDDVYVHNRCEQIIRK-----CKKCGMLPLDLKRLRESKREIELATW 262
OY 221 -----PPDNA--P-----NSFYRALY---P 237
D 263 DKGVGPRSPDASPSGSRKLEDEDAVAVRHGSSLSGSDAGVYKDSDFEFTYRPMK 322
OY 238 KIIODIETIESNMRCGRHSIORHCRSEYK-----GYVCLQYDDOKIYSLRNTIKIMD 293
D 323 EVTKDRFKVGTNNKYG-----GSKTFKHTNGVMCLQFEDNLLATGSDYPTTIKIMD 375
OY 294 KNTLECKRLITGHTGSLVCLQYDERVITITSSSFTVAVRVNVMGEMLTIIHCEAVLH 353
D 376 TENGEBELRLRGHESGRICQFDPDKLISSMRTIKVMMWRKGECLISTYTGIRGVIGL 435
OY 354 RENNMMVTCSSKDRSLAVWDMASPTDITLRLVLRVRAAVN--VDDKYIVASAGDRT 411
D 436 HFDASILASGSDVKTVMYVNFEDKSTFSLR-----GHTDWMVNAVVDTSRTEVSASDDCT 491
OY 412 IKVMNNTSGCFVRLTNGH-----KRGIAAC----- 435
D 492 VRLMDLDTKTCIRTFHGHWGQVOVPLPREFEFEHDAECENDLSTSGDANPESIOA 551
OY 436 -----LQYRD-----RLVYSGSSDNTIRLMDIEGACLRVLEGHEEL 472
D 552 SMGLEPRAVASQSAFSTFSDNGRAAPRYMVSALDSTIRLMEETTIGRLTFPGHLSG 611
OY 473 VRCIRPDNRKIYSGAIDGKTKVMDLVAAALDPRAPAGTLCLRTLVESGRVRLQDEFOI 532
D 612 VMALGADTLRIYSGADRMKIMD-----PRTGKCEPFTGSHGSPVTCIGLDSRF 662
OY 533 VSSSHDITILIMDF 546
D 663 ATGSEDCVAMYSF 676

RESULT 7
MT30_YEAST STANDARD: PRT: 640 AA.
AC P39014;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DN MET30 PROTEIN.
GN MET30 OR YII046M.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=96069360; PubMed=8524217;
RA Thomas D., Kuras L., Barbey R., Cherest H., Blaiseau P.L.,
RA Surdin-Kerjan Y.;
RT "Met30p, a yeast transcriptional inhibitor that responds to S-
RT adenosylmethionine, is an essential protein with WD40 repeats.";
RL Mol. Cell. Biol. 15:6526-6534(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Churchill C.M., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Fraser A.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
CC GENES EXPRESSION.
CC -!- SUBUNIT: SEEMS TO INTERACT WITH MET4.

```

```

CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE MET30/SCONE/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 246861; CAAB6905.1; -
DR EMBL: 126505; AA96717.1; -
DR SGD: S0001308; MET30.
DR InterPro: IPR001800; -
DR InterPro: IPR001810; -
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 6.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PROSITE: PS50181; PROX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Transcription regulation; Methionine biosynthesis;
DR Cysteine biosynthesis; Repeat; WD repeat.
FT DOMAIN 181 227 F-BOX.
FT REPEAT 300 328 WD 1.
FT REPEAT 340 368 WD 2.
FT REPEAT 380 408 WD 3.
FT REPEAT 419 449 WD 4.
FT REPEAT 461 499 WD 5.
FT REPEAT 509 538 WD 6.
FT REPEAT 550 578 WD 7.
FT REPEAT 607 635 WD 8 (POTENTIAL).
FT CONFLICT 61 61 M -> I (IN REF. 1).
SQ SEQUENCE 640 AA; 72835 MM; 5135D4BCA2E1EB97 CRC64;

Query Match 18.0%; Score 545; DB 1; Length 640;
Best Local Similarity 28.9%; Pred. No. 4.8e-33;
Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

OY 34 IPEKSLRQTYNSCARLILN-OETVCLASTAMKTENCYAK-TKLANGTSMIVKORLKS 91
D 86 LPEVNFYKCYVHNNDIDQSPYHNTACYKODLKRQEIYANAKILPILQSDIHHISKYS 145
OY 92 ASYEKEKELCYAFEPQWSESDOVERFENHILISOMCHYOHGHINSLYKRLPMDFTALPAR 151
D 146 NSNDKIRKL-----LIDGILSTSCPPOLSYISLVTNMIKIDFTSLP-- 188
OY 152 GLDHAENILSYLDAKSICAAELVCKEYRVATSDGMLKKLIERAVRTDSLRLGLAER- 210
D 189 --QELSKLILSLYDQSLCNATRVCKRMOKLADDDRVYHNCQIIT-----DKKC 236
OY 211 ---GNGOYLFKNK-----PPDNA--PNSFYRALYPKIIODIETIESNMRCGRHSIQ 258
D 237 RYCGMGDLPLHMKRARIQONSTGSSNADIDQOTRPAKVIYREFKVESMWRG----- 291
OY 259 RIHGRSEYK-----GYVCLQYDDOKIYSGLDNDNTIKIMDKNTLECKRLITGHTSVCLO 314
D 292 --HCRIOEFKGMHGVLLLOEFYRLLFYGSYDSTIGIDLEFGKILRLISGHSQGVKTY 349
OY 315 YDERVITITSSSFTVAVRVNVMGEMLTIIHCEAVLHRLRNNNGMMVTCSSKDRSLAVWDM 374
D 350 FDDRRLITGSLDKTIRVNNYITGECISTYRGHSVLSVDSYQAVIYSGADKTVKWHV 409
OY 375 ASPTITLRLVLRVRAAVNVDFDKYI--VSASGDITKVMWTSICEFVRLTNGH--- 429
D 410 ESRCTYTLR-----GITEVNVCKYKILPKSFSCSDSDITIMMDIRNRSCLKYVRGHVQ 465
OY 430 KRGIACLQYRD--RLVYSGSS----- 448

```


modified and this statement is not removed. Usage by and for commercial CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: 27957; CAB02129.1; -
 DR WormRep: F55812.3; CE16120.
 DR InterPro: IPR001680; -
 DR InterPro: IPR001810; -
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR PROSITE: PS0181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 5.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat: WD repeat.
 FT DOMAIN 113 159 F-BOX.
 FT REPEAT 245 274 WD 1.
 FT REPEAT 286 316 WD 2.
 FT REPEAT 328 356 WD 3.
 FT REPEAT 368 396 WD 4.
 FT REPEAT 408 438 WD 5.
 FT REPEAT 453 481 WD 6.
 FT REPEAT 493 522 WD 7.
 SQ SEQUENCE 579 AA: 64275 MW: 0647245AF5964663 CRC64;

Query Match 17.1%; Score 520; DB 1; Length 579;
 Best Local Similarity 28.8%; Pred. No. 3e-31;
 Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

QY 72 KTKLANGSTSMI-----VPROR-K-LSASYEKEL-----CYKFEQSESDOVE 115
 DB 27 ESSSYNGSSSYNADKLSSSRPLQHKLDLSASPSRNDLPRVHEHLALFKDLSQMD 86
 QY 116 FVEHLISOMCHYOHGHNLSLKLQRFITLAPKGLDIAENILSYLDAKSLCAELY 175
 DB 87 AFTRLQESNMKTNIRQLRAIEHPHOFDLSCLPV---ELGKKILHNLGYDLKVAQY 142
 QY 176 CKERYVTSNGMLMKL-IERMV-----RTDSLWGLAERMGWQYLKKNRPPGNAP 227
 DB 143 SKWKKLISELDKTKMSLGEVEFKHNDPTDRVYGANOGTAIAG-----VITPHIOP 195
 QY 228 PN-SEYRALPKIIODI-----ETIESNMRCGRHSLSQIRKSETSKGYCL 273
 DB 196 CDLVHREFLKLQKFGDIFERADKSRYLRAKLEKNMNPINGSAY-LRGHEDVITCM 254
 QY 274 QYDDQIVSLRNTIKIKMKNLECKRILITGHTGSYLQYDE--RVITGSSDSTVY 331
 DB 255 QIHDVLYVTSNDNTLKWCIIDKGEVWYTLVGTGGMWTSQISQCGRYIVSGSTDRTVKY 314
 QY 332 MDVNTGEMLTLLHHCDAVLHLRPNNGMVTGSKDRSIAVMDASPTDITLRVLVGHRA 391
 DB 315 WSYVDGSLHTLQGHSTVCMAMAGSILVTSRDTTLRWYDVEGRHLA---TLGHHA 371
 QY 392 AVNVVDEDDKYIVASASDRITIKWNTSTCEVFTLNGHKRGACLOYRDR--LVSSGSSD 449
 DB 372 AVYCVQDGTIVYSGDYFTKIMNATGRCIFITLGHNNRIVSLFESRSIVCGSSLD 431
 QY 450 NTRLRMDI---EGACIRVLEGEHEELVRCIRPNKRIYSGAYDKIKWDLVLAALDPRAP 506
 DB 432 TSTRVWDFTRPEQCECVALLQGHSTLSGMQLRGNLIVSCNADSHRVMDI-----H 483
 QY 507 AGTLCLRTLVHSGRVERLO-FDEFQIVSSSHDITLMD 545
 DB 484 EGT-CVAMHLSGHRSAITSLQWFGRRNVAVSSDDGTAKLMD 522

RESULT 10
 ID YAF1_SCHPO STANDARD: PRT: 267 AA.
 AC 009855;
 DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL TRP-ASP REPEATS CONTAINING PROTEIN C29E6.01 IN CHROMOSOME
 DE I (FRAGMENT).
 GN SPAC29E6.01.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Jones L., Murphy L., McNeill A., Simpson I., Harris D., Barrell B.G.,
 RA Rajandream M.A., Walsh S.V.;
 RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: 266525; CAA91423.1; -
 DR InterPro: IPR001680; -
 DR Pfam: PF00400; WD40; 6.
 DR PROSITE: PS00678; WD_REPEATS_1; 2.
 DR PROSITE: PS50082; WD_REPEATS_2; 5.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Hypothetical protein; Repeat; WD repeat.
 FT NON_TER 1 1
 FT REPEAT 20 50 WD 1.
 FT REPEAT 62 90 WD 2.
 FT REPEAT 106 136 WD 3.
 FT REPEAT 148 176 WD 4.
 FT REPEAT 188 216 WD 5.
 SQ SEQUENCE 267 AA: 30324 MW: C76679F6A062C73A CRC64;

Query Match 17.1%; Score 519.5; DB 1; Length 267;
 Best Local Similarity 38.0%; Pred. No. 1.2e-31;
 Matches 104; Conservative 58; Mismatches 83; Indels 29; Gaps 6;

QY 286 DMTIKIMDKNTLECKRILITGHTGSVLCQYDER--VIITGSSDSTVVRWMDVNTGMLNTL 343
 DB 1 DRTVSMDVNSRFLIKYKYGHSGLVCLDFCRRNRLVSGSSDSTITLMDWONRRPLKYV 60
 QY 344 IHHCDAVLHLRPNNGMVTGSKDRSIAVW--DMASPTDITLRRVYVGHRAAVNVVDFDK 401
 DB 61 FGHITNVLVGVSENYIISSRDHTARVRLDATSPAEACM-HVLRGHLSAVNSVQYSK 119
 QY 402 --YIVSASGDRITIKWNTSTCEVFTLNGHKRGACLOYRDLVYVSGSSDNTIRLMDIEC 459
 DB 120 TGLIVTASSDRLTRMDITGTCIRIHAHQIGACQNGKFIVSGSSDLIRIFEASS 179
 QY 460 GACLRVLEGEHEELVRCIRPNKRIYSGAYDKIKWDLVLAALDPRAPAGTLCRLTVH- 518
 DB 180 GLRLRLQGHEDLIRTVRNDKRIYSGGIDGTIVRTW-----FNTGEQHC 224
 QY 519 -----SGRFRLOFDEQIVSSSHDITLMD 546
 DB 225 VLHNSRNSRVFGLQDHRRIITACTHSSETLVWNE 258

RESULT 11
 ID CC4_CANAL STANDARD: PRT: 684 AA.
 AC P53699;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

```

DT 01-OCT-2000 (rel. 40, last annotation update)
DE CELL DIVISION CONTROL PROTEIN 4.
GN CDC4.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Microsporita; Saccharomycetales; Candida.
RN NCBI_TaxID=5476;
RP SEQUENCE FROM N.A.
RC STRAIN=SGY126;
RA Shieh J.C., White A.M., Rosamond J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC VARIOUS ASPECTS OF SPOREATION. REQUIRED FOR HET1-HTB1 LOCUS
CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X96763; CA65538.1; -
DR InterPro: IPR001680; -
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINBPT.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 4.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT DOMAIN 212 258 F-BOX.
FT REPEAT 322 351 WD 1.
FT REPEAT 363 391 WD 2.
FT REPEAT 403 431 WD 3.
FT REPEAT 442 468 WD 4 (POTENTIAL).
FT REPEAT 478 506 WD 5.
FT REPEAT 519 549 WD 6.
FT REPEAT 561 589 WD 7.
SQ SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;

```

```

Query Match 15.7%; Score 477.5; DB 1; Length 684;
Best local similarity 25.4%; Pred. No. 5, 3e-28;
Matches 136; Conservative 95; Mismatches 184; Indels 121; Gaps 16;

```

```

DB 399 HIFGHTSTIRCLDIHPAVIGNKQDQEDIVPEEYLLITGSRDHNHAWKLPVYDSD 458
OY 375 -----ASPTDILR-----RLV 387
DB 459 YIETGEGEPDNPYLAVLSGHQSVASISGYNIITSSYDSIVRWDLDDGCHYVQ 518
OY 388 GH--RAAVNVVDEDDKYIVSASGRTIKYWNSTCEFEVTLNGKRGKIACLRDLV 445
DB 519 GHLDVRYSTALDPHSKTCPSGMSDNIWNVNFETGLKRVLVGHASVGLDLVDVLS 578
OY 446 GSSDNTIRLMDECGACLVLEBHELVRCIRDNKRIVSGAVDGKIKWDLVAALDPPA 505
DB 579 AAADATLRIMDKATGLKRGKLGHAALICFEHDLGVVSGS-EKMLKLMV----- 629
OY 506 PGTLCLFLVHESGVRFLQD-----EQIVSSHDDTL-TWDLNDPPAQ 553
DB 630 EKGTFRDLSDVGTGIGVYREDYKRCVAVQRRIWDEGETTIELDF-SQPLQ 684

RESULT 12
HET1_PODAN STANDARD; PRT; 1356 AA.
AC 000808;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE VEGETABLE INCOMPATIBILITY PROTEIN HET-B-1.
GN HET-B-1.
OS Podospota anserina.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Podospota.
OX NCBI_TaxID=5145;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009891; PubMed=7557402;
RA Saube S., Turc B., Begueret J.;
RT "A gene responsible for vegetative incompatibility in the fungus
RT Podospota anserina encodes a protein with a GTP-binding motif and G
RT beta homologous domain."
RL Gene 162:135-139(1995).
CC -1- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
CC HET-C.
CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L28125; AAA5775.1; -
DR InterPro: IPR001680; -
DR Pfam: PF00400; WD40; 10.
DR PRINTS: PR00320; GPROTEINBPT.
DR PROSITE: PS00678; WD_REPEATS_1; 10.
DR PROSITE: PS50082; WD_REPEATS_2; 10.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW GTP-binding; Repeat; WD repeat.
FT NP_BIND 300 307 GTP (POTENTIAL).
FT REPEAT 839 869 WD 1.
FT REPEAT 881 911 WD 2.
FT REPEAT 923 953 WD 3.
FT REPEAT 965 995 WD 4.
FT REPEAT 1007 1037 WD 5.
FT REPEAT 1049 1079 WD 6.
FT REPEAT 1091 1121 WD 7.
FT REPEAT 1133 1163 WD 8.
FT REPEAT 1175 1205 WD 9.
FT REPEAT 1217 1247 WD 10.
SQ SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;

```

Query Match 15.0%; Score 455.5; DB 1; Length 1356;
 Best Local Similarity 26.0%; Pred. No. 5.5e-26;
 Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

QY 28 EPPRKIIPEKNSLRQYNSCAR-ICLNOETVCLASTAMK-----TENCVAKTKLANGTSS 81
 DB 581 EMPDVSDEDSLEIEIKGSEFLIERIRYEVHQAKDFLGTASDKSNASQCAFE 640
 QY 82 MIYK-----QRLASYEKE--LCYKFEQMSQVEVEHL--ISQCHY 127
 DB 641 LVFPTGEDVSYIIFWMSLVMSOKLRDIYCLNAPGFLIDNVRPPDLATVRYSCIY 700
 QY 128 QHGINSYLRP-----MLORD-----FIT-----ALPARGLDHI--AEN 159
 DB 701 WIDLRLVLSSTSKWYHLLDDODDHRFLTTRYLWLEALSRLALP-EGINAIROLES 759
 QY 160 ILAYLDAKSLCAELVCKEYRVTSDGMLKKLIER--MYRDSLWRGLAERRGQYL 216
 DB 760 LIGHTRIRLIA--IVRDGYR--ALSYRMIEKAPLQATYSAL-----V 800
 QY 217 FKNPPPGNAPNSFYRALY--PKIIDIEETESMWRGGRSL----- 257
 DB 801 F-----APTDSMIKKIRKEEPGNIISTISVEAEAMNACTOTLEGHSSVLSVAFSAD 852
 QY 258 -QRIHRS-----ETSKG-----VYCLQY--DDOKIVSGLRDNTIKIMDK 294
 DB 853 GQRAVSSGDKTIKIMDTASTGTQTLLEGHGSVMSVAFSPDERVAVSGSDDTIKIMDA 912
 QY 295 NTLCKRILFGHGSVLCLOY--DERVIITGSSDSIVRVMDVTGEMLNTLIHHCVAVLH 352
 DB 913 ASGCTQTLLEGHGRVQSVAFSPDQRAVSSGSDHTIKIMDAAGTCTQTLLEGHGSVLS 972
 QY 353 LREN-NGMAYTC-SKDSIAVMDASPTDITLRVIVGHAAVAVDF--DQIYIASAG 408
 DB 973 VAFSPDQRAVSSGSDHTIKIMDTASTGTQTLLEGHGSVMSVAFSPDQRAVSSGSD 1029
 QY 409 DRTIKVNTSTGEFVFTLNGHKRQIACLOY--DRLVVGSSSDNTILMDIEGACLRVL 466
 DB 1030 DKTIKIMDTASTGTQTLLEGHGSVMSVAFSPDQRAVSSGSDHTIKIMDAVSGTCTQTL 1089
 QY 467 EGHELVRCIRF--DNRIIVSGAYDGRIKIMDLVLAALDPRAPAGTCLRLTVEHSGRFR 524
 DB 1090 EGHDSVMSVAFSPDQRAVSSGSDHTIKIMD-----AASGT-CTQTLLEGHGVHVS 1140
 QY 525 LQF--DEFOIVSSHDTITLIMPLNDPAAQ 553
 DB 1141 VAFSPDQRAVSSGSDHTIKIMDAAGTCTQ 1171

RESULT 13
 POP1_SCHPO STANDARD; PRT; 775 AA.
 AC P87060;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE WD-REPEAT PROTEIN POP1.
 OS POP1 OR SPBC262.18.
 GN Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97347242; PubMed=9203581;
 RA Komlami K., Toda T.;
 RT "Fission yeast WD-repeat protein pop1 regulates genome ploidy through
 ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1
 and the S-phase initiator Cdc18.";

RL Genes Dev. 11:1548-1560(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Rajadream M.A., Barrell B.G., Devlin K., Churcher C.M.;
 RL Submitted (Mar-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN MAINTENANCE OF PLOIDY THROUGH PROTEASOME
 CC DEPENDENT DEGRADATION OF CDK INHIBITOR RUM1 AND S-PHASE INITIATOR
 CC CDC18. FUNCTIONS AS A RECOGNITION FACTOR FOR RUM1 AND CDC18, WHICH
 CC ARE SUBSEQUENTLY UBIQUITINATED AND TARGETED TO THE 26S PROTEASOME
 CC FOR DEGRADATION.
 CC -1- SUBUNIT: BINDS TO CDC18.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC
 DR EMBL: Y08391; CAA69671.1; -
 DR EMBL: AL022103; CAA17898.1; -
 DR InterPro: IPR001680; -
 DR InterPro: IPR001810; -
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 6.
 DR PRINTS: PR00320; GPROTEINBPT.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 3.
 DR PROSITE: PS50082; WD_REPEATS_2; 4.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 298 345 F-BOX.
 FT REPEAT 444 472 WD 1.
 FT REPEAT 484 538 WD 2.
 FT REPEAT 575 603 WD 3.
 FT REPEAT 615 645 WD 4.
 FT REPEAT 657 687 WD 5.
 SQ SEQUENCE 775 AA; 87816 MW; B06EDBA6553EBC1 CRC64;

Query Match 14.9%; Score 453; DB 1; Length 775;
 Best Local Similarity 25.8%; Pred. No. 4.1e-26;
 Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;

QY 9 QEKALKFNSSSEREDC-----NNGEPPRKIIPEKNSLRQY-NSCARICLNOET 56
 DB 121 REKCLRRNSSLSNLHANKRFLFNSQSDGNKKNETFPSTNSVVFYNNCDSEVASET 180
 QY 57 VC-----LASTAMKTENCVAKTKLANGTSMIYPRQRKLSA----- 92
 DB 181 TFSLDAPNNSVNYSYFSPNLLGNDSKTQSPFPHSSSSSHNSLHPVLIYDPSSENPISHP 240
 QY 93 -----SYEKKEKL-----CVAYFEQMSQVEVEHLISQCHYQHGINSYLRKPLMDQRF 144
 DB 241 SNHLSQKNAVLKQLQSSFEKRLPESVRQYLLPHLLSRGCHNAYONHKLILRPFGQNF 300
 QY 145 ITALPARGLDIAENILSYLDAKSLCAELVCKEYR-VTSDGMLKKL-----I 193
 DB 301 LINGFPA-----FITNLVLTLDAPSLCAVQSVSHHYKLVSSNEELMKSILFLKDGFMDSI 356
 QY 194 EKNVPTDSLWRGLAERRGQYLFRKKNPPDGNAPNSYRALYRKIIDIEETESMWRGSG 253
 DB 357 DSKIRTMCLQSLA-----CATMKRVYFRHFNLRERW----- 389
 QY 254 RSLQRIHCRSTSKGVYC---LOYDQKIVSGLRDNTIKIMDKNTLCKRILFGHTGTSV 310
 DB 390 LIAPEKIRKCSFPHGVRLITKLOFDGDKLIIVTSPRINITYDFRTVGLISLSEHEDV 449
 QY 311 LCLQYDERVIITGSSDSIVRVMDVTGEMLNTLIHHCVAVLHLRENNGMAYTCSKNSIA 370

```

DB 450 WFEVYGDITLVGSDRTVWDLRTGECKQVYFGTSTIRCKIYQGNQSTDTD----- 505
OY 371 WMDMSPDITLRVLVGHRAAVNVDPDCKIVASASGRITKVMN----- 416
DB 506 -----DVEKE-----NRPASNDANSMPYIISRRCTIRLWMLPCLDDPPVYVNE 552
OY 417 -----TSTCEVFTLNGHK-----RGIACTLOYRDLVYVSSSDNTIRLMDIECGAC 462
DB 553 NPDQNDNFTSATTNPEYITRLGHNTDSVREAVCL---GLIYASASYDGLRLVWKASGTVC 609
OY 463 LKVLBSHELKVCIRFDNKR---IVSGAIDGKTVMVDLVALDPRAPAGTLCRLTVEHSG 520
DB 610 LHVLRKHVRVYSVYIINPSROOCISAGTDKIRIWNL-----ESGEL-LQTLIGHSN 660
OY 521 RVERLOFDEFQIVSSS---HDDTILIMD 545
DB 661 LVSQVTFNONIIVSASAPDITSLRWMD 687

RESULT 14
CC4_YEAST STANDARD: PRT: 779 AA.
AC P07834;
DT 01-AUG-1988 (rel. 08, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE CEL4 DIVISION CONTROL PROTEIN 4.
GN CDC4 OR YFL009W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86011240; PUBMED=3309335;
RA Yeohem J., Byers B.;
RT "Structural comparison of the yeast cell division cycle gene CDC4 and
RT a related pseudogene.";
RL J. Mol. Biol. 195;233-245(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=93400292; PUBMED=7670463;
RA Murkamt Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasunuma S.-I., Sasunuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 1-579 FROM N.A.
RX STRAIN=S288C / AB972;
RA Barrell B.G., Churcher C., Rajandream M.A.;
RT Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
RL REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
RL POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
RL DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
RL VARIOUS ASPECTS OF SPOULATION. REQUIRED FOR HTAI-HTBI LOCUS
RL TRANSCRIPTION ACTIVATION.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL, X05625; CAA29113.1;

```

```

DR EMBL, D50617; BAA09229.1; -
DR EMBL, D31600; BAA06495.1; -
DR EMBL, 246255; CAA86341.1; -
DR PIR, A26867; A26867.
DR SGD, S0001865; CDC4.
DR InterPro; IPR001680; -
DR InterPro; IPR001810; -
DR Pfam; PF00646; F-box_1.
DR Pfam; PF00400; WD40_6.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS50181; FBOX_1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KM Cell division: Mitosis; Sporulation; Repeat; WD repeat.
FT DOMAIN 272 319 F-BOX.
FT REPEAT 380 408 WD 1.
FT REPEAT 420 449 WD 2.
FT REPEAT 461 493 WD 3.
FT REPEAT 528 556 WD 4.
FT REPEAT 568 598 WD 5.
FT REPEAT 630 658 WD 6.
FT REPEAT 669 698 WD 7.
FT CONFLICT 460 460 K -> E (IN REF. 1).
SQ SEQUENCE 779 AA; 86089 MW; 0348F2F8FA78F3BC CRC64;

```

Query Match 13.2%; Score 399; DB 1; Length 779;

Best Local Similarity 24.3%; Pred. No. 4,1e-22; Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

```

OY 56 TWCLASTAMKTCNVAKTRLANGTSSMIVPKOKRISASYEKER-----ELCVKFEQW 108
DB 194 TPLPLAKTKTINN-----NNNIADLESKDSIISPEYLSDFSAIINNMLPHAYFK-- 244
OY 109 SESDQVEFEHLISQCHQHGHINSYLKPMQORDITLALPARGLHLENISYDAKS 168
DB 245 -----NILEFLVANNDRSELSDLGLIKDNLKRDITSLP-----EISLKFNYLQFED 294
OY 169 LCAAEIVCKEWY-VTSDMLMKKLI--BRMYRTDSLMLGLAERRGQGYLFRKKNPPDGN 225
DB 295 IINSLGVSONMKKIIRKKSIMKKLISNFV-----SPKGF 331
OY 226 APPNSFYRALYPTIID-----IEF--TESNWRGGRSLORICRSETSGVYCLQYD 276
DB 332 NSLNLKLSQKYPRLSQODRLSLFLENIFLKNWYNPKFVPOFTLRGHMTSVITCLOPE 391
OY 277 DQKIVSGLADNFKIKMDKNTLECKRLITLGHGTVCLQYDE-RVITTGSSDSIVRVWDVN 335
DB 392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGVMALKYAHGGLIVSGSDTRIVRWDIK 451
OY 336 TGEMLNTLIHCEAVLHLRFNNQMMVTGSKDRSIAVMDASPTDITLRVLVGHRAAVNV 395
DB 452 KG-----CCT-----HYFGHNSIVK 468
OY 396 VFEDD---KYIVSASGDRITKVMN-----STCE---FVFTLNG 428
DB 469 LDIVEKKNKIYITGSRDNLHAWKLPKSSVDPDHEEHDPYLVFTPRENPYFVCLNG 528
OY 429 HKRGICLOYRDLVYVSSSDNTIRLMDIECGACLVLSGHELVLCIRFDN--KIVSG 486
DB 529 HNASVTVSGHGNIIVSGSYDNTLIWMDVAQMKCLYILSGHTRISTYIDHKKRCISA 588
OY 487 AVDGKIKVMDL-----VALDPRAPAGT--CLRTLVHSGVFRLOPDEPQIVSS 535
DB 589 SMTTTRIMDLLENINWNGECSTATNSAPCAKILGMYTLQGHITALVGLRLSDKFLVSA 648
OY 536 SHDDTILIMDLNDPAAQAEPPSPSRITYY 566
DB 649 AADGSIRGMD-AND-----YSRKFSY 668

```

RESULT 15

POP2_SCHPO STANDARD: PRT; 703 AA.
 ID POP2_SCHPO 014170;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE WP-REPEAT PROTEIN POP2 (PROTEOLYSIS FACTOR SUD1).
 GN POP2 OR SUD1 OR SPAC4D7.03.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 XX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SFRAIN-972;
 RA Wolf D.A., Jackson P.K.;
 RL submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98318628; PubMed=9653157;
 RA Jallepalli P.V., Tien D., Kelly T.J.;
 RT "sudd" targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1
 RT proteins for degradation and stops unwanted diploidization in fission
 RT yeast.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SFRAIN-972;
 RA Gentiles S., Churcher C.M., Bartell B.G., Rajandream M.A., Wood V.;
 RL submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PUTATIVE COMPONENT OF E3 (PROTEIN-UBIQUITIN LIGASE)
 CC COMPLEX: TARGETS CYCLIN-DEPENDENT KINASE-PHOSPHORYLATED CDC18 AND
 CC RUM1 PROTEINS FOR DEGRADATION.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
 CC
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF038867; AAB95480.1; -;
 CC EMBL: AF064515; AAC39496.1; -;
 CC EMBL: Z98602; CAB1275.1; -;
 CC InterPro: IPR001880; -;
 CC InterPro: IPR001810; -;
 CC Pfam: PF00646; F-box; 1.
 CC Pfam: PF00400; WD40; 6.
 CC PRINTS: PR00320; GPROTEINRPT.
 CC PROSITE: PSS0181; FBOX; 1.
 CC PROSITE: PSS00678; WD_REPEATS_1; 3.
 CC PROSITE: PSS0082; WD_REPEATS_2; 6.
 CC PROSITE: PSS0294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 236 F-BOX.
 FT REPEAT 389 417 WD 1.
 FT REPEAT 429 473 WD 2.
 FT REPEAT 505 533 WD 3.
 FT REPEAT 545 575 WD 4.
 FT REPEAT 587 615 WD 5.
 FT REPEAT 625 654 WD 6.
 SO SEQUENCE 703 AA; 79605 MW; 0336A0568C152396 CRC64;

Query Match 13.1%; Score 396.5; DB 1; Length 703;
 Best Local Similarity 23.3%; Pred. No. 5,6e-22;
 Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

QY 14 KFMNSERED--CNGGPEPRKI-----IPEKNSLRQTYNSCARLCL---NQETVCL 59

Db 96 KFMNSIENVSCKLDHSPDPVSVDGFSISLVQPQNFVLYSHSLPPKILSIDRRNRIRKLD 154
 QY 60 ASIAMTENCVAKTAKLIANTSSMIYPKOKLSASVEK-EKELCVAYFEQWSSDQVEEVE 118
 Db 155 NTSISNSDNFPPSPKV--DTSNTVSPGSKPISDEEDLNLOSIVQTFDDLPGLQ-SYAF 211
 QY 119 HITSQCHYQHGH-INSYLPKMLQDFFIALPARGLDIAENILSYLAKSICAAELVCK 177
 Db 212 FQLLRSCNQSRLNLECEPILKDLISNLPF-----SIVGSILNLNDIHSLSRLVSP 267
 QY 178 EWMYV---TSDGMMLKKLIERMVRTSLMGLAERCGOYLFRNKKPPDGNAF---N 229
 Db 268 TWMRLDVHTS--YWKHMF-----SLFGQINENDW---KYAMPNLRRPPLFAND 312
 QY 230 SFYRALYPRKIIDIETIESNMRCGRHSLOF----- 259
 Db 313 QISDDYFPEIFK-----RHFLNRKRLMPLPSIPPSHLSPPIHPNEMITSLILH 360
 QY 260 -----IHCRSETSK-GVYCLQDQKIVSGLRDNTIKIMDKNT 296
 Db 361 KDRITTSGGTITQIHNAITGVLEARLEBKRGVAVAKIHENTLVSGSIDKTVRWNTIEK 420
 QY 297 LECKRLTGTSGSVLCLOY-----DERVIITGSSDSTVYRWVDVTGEML 340
 Db 421 AKCTHIFRGHISIRCLTELIVPSRLIRHGVETVEPDQPIYVSGSRDHLRWV----- 472
 QY 341 NTLIHCEAVLHFRFNGMWTCSKDRSIAYVMDASPTDITLRYLVGHRRAVNVDEDD 400
 Db 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNPYFV--HTLIGHTSVRTISGIG 517
 QY 401 KYIVSASGDRITKIVNTSYCEFRVLNGHKGRIACLOYRD--RLVVGSSDNTIRLMDIE 458
 Db 518 DILVSGSYDSIRIMRVSGECLYHLRHSLSRIYSVLVEPEENICISMSKSIWMLDS 577
 QY 459 CGACLRVLEGHEELVRCIFPDNKRIVSGAYDGKIKVMDLVAAIDPRAAGTLCRTLVEH 518
 Db 578 TGTCKRYVLEGHDAFVYLLNVFQNRILISGASDSTIRIMD---LNTGKP-----LNVLP 628
 QY 519 SGRVFLQDFDEFOIVSSSHDITLIMD 545
 Db 629 SGYISFVSDERKII-SGMDGSVKLMD 654

Search completed: June 7, 2001, 10:38:26
 Job time: 145 sec

THIS PAGE BLANK (USPTO)

Thu Jun 7 10:44:15 2001

us-09-415-795-4.rsp

Page 13

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 10:35:02 ; Search time 33.7 Seconds
(without alignments)
1978.970 Million cell updates/sec

Title: US-09-415-795-4

Perfect score: 3034
Sequence: 1 MDPAEAVLQEKALKFNNSS.....PAAQAEPPRSPSRITYISR 569

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_15:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3034	100.0	569	4 Q9Y213	Q9Y213 homo sapien
2	3006	99.1	605	4 Q9Y297	Q9Y297 homo sapien
3	2997	98.8	569	11 Q9Q015	Q9Q015 mus musculu
4	2990	98.5	569	11 Q9Z159	Q9Z159 mus musculu
5	2978	98.2	569	11 Q9R1G7	Q9R1G7 mus musculu
6	2400	79.1	529	4 Q9P2S8	Q9P2S8 homo sapien
7	2384.5	78.6	542	4 Q9UKB1	Q9UKB1 homo sapien
8	2384.5	78.6	550	4 Q9Y4C6	Q9Y4C6 homo sapien
9	2358	77.7	508	4 Q9P2S9	Q9P2S9 homo sapien
10	2089.5	68.9	510	5 Q44382	Q44382 drosophila
11	2089.5	68.9	510	5 Q9YDE3	Q9YDE3 drosophila
12	690	22.7	506	3 Q9P7V1	Q9P7V1 schizosach
13	661	21.8	1326	3 Q9VZF4	Q9VZF4 drosophila
14	640	21.1	553	4 Q9NUX6	Q9NUX6 homo sapien
15	520	17.1	587	5 Q44083	Q44083 caenorhabdi
16	365.5	12.0	942	5 Q96611	Q96611 dictyosteli
17	356	11.7	410	13 Q9PTR5	Q9PTR5 gallus galli
18	344	11.3	411	5 Q96698	Q96698 drosophila
19	339.5	11.2	777	3 Q9USN3	Q9USN3 schizosach

20	334.5	11.0	422	4 Q9UKT8	Q9UKT8 homo sapien
21	334.5	11.0	422	4 Q9UKA3	Q9UKA3 homo sapien
22	332.5	11.0	317	10 Q9M222	Q9M222 arabidopsis
23	332	10.9	454	11 Q9QUH1	Q9QUH1 mus musculu
24	330	10.9	333	10 Q9SY00	Q9SY00 arabidopsis
25	325	10.7	277	11 Q9R2A6	Q9R2A6 mus musculu
26	325	10.7	391	11 Q35392	Q35392 mus musculu
27	324.5	10.7	334	4 Q9NMW7	Q9NMW7 homo sapien
28	324.5	10.7	362	4 Q9UGP9	Q9UGP9 homo sapien
29	313.5	10.3	481	5 Q9VPR4	Q9VPR4 drosophila
30	313.5	10.3	485	4 Q9NVX2	Q9NVX2 homo sapien
31	312.5	10.3	330	4 Q9NUL4	Q9NUL4 homo sapien
32	312	10.3	876	10 Q9LFE2	Q9LFE2 arabidopsis
33	308	10.2	480	5 Q96995	Q96995 drosophila
34	304.5	10.0	1205	4 Q9UJ66	Q9UJ66 homo sapien
35	304.5	10.0	1205	4 Q9UJ59	Q9UJ59 homo sapien
36	301.5	9.9	476	13 Q93531	Q93531 xenopus lae
37	299.5	9.9	1205	4 Q9UJ65	Q9UJ65 homo sapien
38	298.5	9.8	1171	4 Q9UJ58	Q9UJ58 homo sapien
39	298.5	9.8	1205	4 Q9UJ67	Q9UJ67 homo sapien
40	298.5	9.8	1205	4 Q9UJ64	Q9UJ64 homo sapien
41	298.5	9.8	1205	4 Q9UJ63	Q9UJ63 homo sapien
42	297	9.8	502	3 Q74855	Q74855 schizosach
43	295.5	9.7	1205	4 Q9UJ60	Q9UJ60 homo sapien
44	293.5	9.7	1205	4 Q9UJ62	Q9UJ62 homo sapien
45	293	9.7	760	10 Q9LHNS	Q9LHNS arabidopsis

ALIGNMENTS

RESULT	ID	Q9Y213	PRELIMINARY:	PRT:	569 AA.
AC	Q9Y213				
DT	01-NOV-1999	(TREMBLrel. 12, Created)			
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)			
DE	01-JUN-2000	(TREMBLrel. 14, Last annotation update)			
DE	BETA-TRANSDUCIN REPEATS CONTAINING PROTEIN.				
CN	BETA-TRCP OR FBW1A.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LYMPHOID;				
RX	MEDLINE=98325370; PubMed=9660940;				
RA	Margolin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,				
RA	Thomas D., Strebel K., Benarous R.;				
RT	"A novel human WD protein, h-beta Trcp, that interacts with HIV-1 Vpu				
RT	connects CD4 to the ER degradation pathway through an F-box motif.";				
RL	Mol. Cell 11:565-574(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Cenciarelli C., Chlaui D.S., Guardavaccaro D., Parks W., Vidal M.,				
RA	Pagano M.;				
RT	"Identification of a family of human F-box proteins.";				
RL	Curr. Biol. 9:1177-1179(1999).				
DR	EMBL: Y14153; CAA74572.1; -.				
DR	EMBL: AF129530; AAF04464.1; -.				
DR	INTERPRO: IPR001800; -.				
DR	INTERPRO: IPR001810; -.				
DR	PFAM: PF00400; WD40; 7.				
DR	PFAM: PF00646; F-box; 1.				
DR	PRINTS: PR00320; GPROTEINRPT.				
DR	PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.				
KW	WD repeat.				
SQ	SEQUENCE 569 AA: 65049 MW: 6E7C4FE7BB520BA CRC64:				
Query Match	100.0%;	Score 3034;	DB 4;	Length 569;	
Best Local Similarity	100.0%;	Pred. NO. 2.7e-244;			

Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLQEKALFKFMSSEEDCNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLA 60
 Db 1 MDPAEAVLQEKALFKFMSSEEDCNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLA 60

QY 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFEQWSESDQVEVEHL 120
 Db 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFEQWSESDQVEVEHL 120

QY 121 ISOMCHYOHGHINSYKPMLOQDFTALPARGLDHAENILSYLDAKSLCAAEVCKEWMY 180
 Db 121 ISOMCHYOHGHINSYKPMLOQDFTALPARGLDHAENILSYLDAKSLCAAEVCKEWMY 180

QY 181 RVTSDGMLMKKLIERNVFTSILMGLAERGMGOYLFEKNKPPDGNAPNSFYRALYPKII 240
 Db 181 RVTSDGMLMKKLIERNVFTSILMGLAERGMGOYLFEKNKPPDGNAPNSFYRALYPKII 240

QY 241 ODIEIESNMRCGRHSIQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
 Db 241 ODIEIESNMRCGRHSIQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300

QY 301 RILFGHGSVYLCLOYDERVITITGSDSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 Db 301 RILFGHGSVYLCLOYDERVITITGSDSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360

QY 361 VTCSKDRSIAMWMASTPDTILRLRVLGHRAAVNVDFDQKIYVSGDRTIKVWNTSTC 420
 Db 361 VTCSKDRSIAMWMASTPDTILRLRVLGHRAAVNVDFDQKIYVSGDRTIKVWNTSTC 420

QY 421 EFVFTLNGHKRGIACTLOYRDLRVVSSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDN 480
 Db 421 EFVFTLNGHKRGIACTLOYRDLRVVSSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDN 480

QY 481 KRIYSGAYDCKIKVMDVLAALDPPAPAGTICLRTLVHSGRVRLQDDEFQIYSSHDDT 540
 Db 481 KRIYSGAYDCKIKVMDVLAALDPPAPAGTICLRTLVHSGRVRLQDDEFQIYSSHDDT 540

QY 541 ILIMDFLNDPAQAEPSPSPRTTYTISR 569
 Db 541 ILIMDFLNDPAQAEPSPSPRTTYTISR 569

RESULT 2

ID 09Y297 PRELIMINARY; PRT: 605 AA.

AC 09Y297;

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE B-TRCP VARIANT E3RS-IKAPPAB (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catartini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RP MEDLINE=99075339; Pubmed=9859996;

RA Yaron A., Hatzudai A., Davis M., Lavon I., Amit S., Manning A.M., Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.; "Identification of the receptor component of the IkappaBalpha-ubiquitin ligase.";

RT Nature 396:590-594(1998).

RL EMBL: AF101784; AAD08702.1; .

RN INTERPRO: IPR001680; .

DR INTERPRO: IPR001810; .

DR PFM: PF00400; WD40; 7.

DR PFM: PF00646; F-box; 1.

DR PRINTS: PR00320; GPROTEINRPT.

DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.

FT NON_TER 1 1

SQ SEQUENCE 605 AA; 68866 MM; 4C67F3B7E400FD37 CRC64;

Query Match 99.1%; Score 3006; DB 4; Length 605;
 Best Local Similarity 94.0%; Pred. No. 6.3e-242;
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDPAEAVLQEKALFKF-----NSSEREDC 24
 Db 1 MDPAEAVLQEKALFKFMSMPRSIMAGCSSIADSMPSRLCLVNPGTGALTARONSEREDC 60

QY 25 NNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLASAMTENCVAKTKLANGTSSMIV 84
 Db 61 NNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLASAMTENCVAKTKLANGTSSMIV 120

QY 85 PKORKLSASYEKEKELCVKFEQWSESDQVEVEHLISOMCHYOHGHINSYKPMLOQDFT 144
 Db 121 PKORKLSASYEKEKELCVKFEQWSESDQVEVEHLISOMCHYOHGHINSYKPMLOQDFT 180

QY 145 ITALPARGLDHAENILSYLDAKSLCAAEVCKEWMYRVTSOGMLMKKLIERNVFTSILMR 204
 Db 181 ITALPARGLDHAENILSYLDAKSLCAAEVCKEWMYRVTSOGMLMKKLIERNVFTSILMR 240

QY 205 GLAERGMGOYLFEKNKPPDGNAPNSFYRALYPKIIDIEIESNMRCGRHSIQRIHCRS 264
 Db 241 GLAERGMGOYLFEKNKPPDGNAPNSFYRALYPKIIDIEIESNMRCGRHSIQRIHCRS 300

QY 265 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRILTGHTGSVLCLOYDERVIITGS 324
 Db 301 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRILTGHTGSVLCLOYDERVIITGS 360

QY 325 SPSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSKDRSIAMWMASTPDTILRLR 384
 Db 361 SPSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSKDRSIAMWMASTPDTILRLR 420

QY 385 VLVGHRAAVNVDFDQKIYVSGDRTIKVWNTSCVEVFTLNGHKRGIACTLOYRDLRVV 444
 Db 421 VLVGHRAAVNVDFDQKIYVSGDRTIKVWNTSCVEVFTLNGHKRGIACTLOYRDLRVV 480

QY 445 SSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDNKRIVSGAYDGIKVMVLAALDPR 504
 Db 481 SSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDNKRIVSGAYDGIKVMVLAALDPR 540

QY 505 APAGTICLRTLVHSGRVRLQDDEFQIYSSHDDTILIMDFLNDPAQAEPSPSPRTY 564
 Db 541 APAGTICLRTLVHSGRVRLQDDEFQIYSSHDDTILIMDFLNDPAQAEPSPSPRTY 600

QY 565 TYISR 569
 Db 601 TYISR 605

RESULT 3

ID 09Q015 PRELIMINARY; PRT: 569 AA.

AC 09Q015;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE UBQUITIN LIASE FMD1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP MEDLINE=99199275; Pubmed=10097128;

RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M., Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A., Nakayama K.-I.;

RT "Ubiquitin ligase Skp1/Cul 1/F-box protein FMD1.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE-99075339; PubMed-9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase";
RL Nature 396:590-594(1998).
DR EMBL; AF081887; AAD17755.1; -;
DR EMBL; AF099932; AAD08701.1; -;
DR INTERPRO: IPR001680; -;
DR INTERPRO: IPR001810; -;
DR PFAM: PF00400; WD40; 7;
DR PFAM: PF00646; F-box; 1;
DR PRINTS: PR00320; GPROTEINRPT.
DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.
KW Ligase.
SQ SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Query Match 98.8%; Score 2997; DB 11; Length 569;
Best Local Similarity 98.6%; Pred. No. 3.2e-241;
Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAEAVLQKALKFNNSSREDCCNNGEPPRKTIIPKNSLRQTYNSCARCLNQETVCLA 60
DB 1 MDPAEAVLQKALKFNNSSREDCCNNGEPPRKTIIPKNSLRQTYNSCARCLNQETVCLT 60
QY 61 STAMKTENCYAKTKLANGTSSMIVPKOKLSASYEKEKELCVYFFQWSSDQVEVEYEH 120
DB 61 STAMKTENCYAKTKLANGTSSMIVPKOKLSASYEKEKELCVYFFQWSSDQVEVEYEH 120
QY 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
DB 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
QY 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQGYLFKNKPPDENAPPNSEFYALVPKII 240
DB 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQGYLFKNKPPDENAPPNSEFYALVPKII 240
QY 241 QDIETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
DB 241 QDIETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
QY 301 RIITGHTGVSVCLOQYDERVITIGSSDSTYKVMVDNTEGEMINTLIHHC EAVLHLRFNNGM 360
DB 301 RIITGHTGVSVCLOQYDERVITIGSSDSTYKVMVDNTEGEMINTLIHHC EAVLHLRFNNGM 360
QY 361 VTCSKRSIAVMDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
DB 361 VTCSKRSIAVMDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
QY 421 EFVRTINGHKGRIACLOYRDRLVYSSSDNTIRLMDIEGACLRVLEGHELYRCIRFDN 480
DB 421 EFVRTINGHKGRIACLOYRDRLVYSSSDNTIRLMDIEGACLRVLEGHELYRCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVLAALDPAPAGTLCRLTVEHSGRVERLOFDEFQIVSSSHDT 540
DB 481 KRIVSGAYDGKIKVMDLVLAALDPAPAGTLCRLTVEHSGRVERLOFDEFQIVSSSHDT 540
QY 541 ILIMDFLNDPAQAEPSPSRRTTYTISR 569
DB 541 ILIMDFLNDPAQAEPSPSRRTTYTISR 569

RESULT 4
Q92159
ID 092159 PRELIMINARY; PRT; 569 AA.
AC 092159;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
ON (1)
RP SEQUENCE FROM N.A.
RA MEDLINE-99145465; PubMed-9990853;
RA Spencer E., Jiang J., Chen Z.J.,
RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT S1mb/Delta-Trcp";
RL Genes Dev. 13:284-294(1999).
DR EMBL; AF112979; AAD04181.1; -;
DR INTERPRO: IPR001680; -;
DR INTERPRO: IPR001810; -;
DR PFAM: PF00400; WD40; 7;
DR PFAM: PF00646; F-box; 1;
DR PRINTS: PR00320; GPROTEINRPT.
DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_6.
SQ SEQUENCE 569 AA; 65047 MW; BC7C7AA44815BED96 CRC64;

Query Match 98.5%; Score 2990; DB 11; Length 569;
Best Local Similarity 98.4%; Pred. No. 1.2e-240;
Matches 560; Conservative 3; Mismatches .6; Indels 0; Gaps 0;

QY 1 MDPAEAVLQKALKFNNSSREDCCNNGEPPRKTIIPKNSLRQTYNSCARCLNQETVCLA 60
DB 1 MDPAEAVLQKALKFNNSSREDCCNNGEPPRKTIIPKNSLRQTYNSCARCLNQETVCLT 60
QY 61 STAMKTENCYAKTKLANGTSSMIVPKOKLSASYEKEKELCVYFFQWSSDQVEVEYEH 120
DB 61 STAMKTENCYAKTKLANGTSSMIVPKOKLSASYEKEKELCVYFFQWSSDQVEVEYEH 120
QY 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
DB 121 ISOMCHYOHGHINSYLYKPMLOQDFITALPARGLDHAENILSYLDKSLCAELVCKEWY 180
QY 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQGYLFKNKPPDENAPPNSEFYALVPKII 240
DB 181 RYVSDGMLMKKLLERAVRTDLSMRGLAERGMQGYLFKNKPPDENAPPNSEFYALVPKII 240
QY 241 QDIETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
DB 241 QDIETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300
QY 301 RIITGHTGVSVCLOQYDERVITIGSSDSTYKVMVDNTEGEMINTLIHHC EAVLHLRFNNGM 360
DB 301 RIITGHTGVSVCLOQYDERVITIGSSDSTYKVMVDNTEGEMINTLIHHC EAVLHLRFNNGM 360
QY 361 VTCSKRSIAVMDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
DB 361 VTCSKRSIAVMDMASPTDITLRVLYGHRAAVNVDFDQKIVYSAGDRTIKVMNTSTC 420
QY 421 EFVRTINGHKGRIACLOYRDRLVYSSSDNTIRLMDIEGACLRVLEGHELYRCIRFDN 480
DB 421 EFVRTINGHKGRIACLOYRDRLVYSSSDNTIRLMDIEGACLRVLEGHELYRCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVLAALDPAPAGTLCRLTVEHSGRVERLOFDEFQIVSSSHDT 540
DB 481 KRIVSGAYDGKIKVMDLVLAALDPAPAGTLCRLTVEHSGRVERLOFDEFQIVSSSHDT 540
QY 541 ILIMDFLNDPAQAEPSPSRRTTYTISR 569
DB 541 ILIMDFLNDPAQAEPSPSRRTTYTISR 569

RESULT 5
Q9R1G7
ID 09R1G7 PRELIMINARY; PRT; 569 AA.
AC 09R1G7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Winston J., Elledge S.J., Harper J.W.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF110396; AAD41025.1; -
DR INTERPRO: IPR001680; -
DR INTERPRO: IPR001810; -
DR PFAM: PF004400; WD40; 7
DR PFAM: PF00646; F-box; 1
DR PRINTS: PR00320; GPROTEINRPT.
DR PROSITE: PS00678; WD_REPEATS;
SQ SEQUENCE 569 AA; 65209 MM; EGGDCAD28D51D9D CRC64;
Query Match 98.2%; Score 2978; DB 11; Length 569;
Best Local Similarity 97.9%; Pred. No. 1.2e-239;
Matches 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 1 MDPAEAVLQEKALFPMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCTINQETVCLA 60
DB 1 MDPAEAVLQEKALFPMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCTINQETVCLT 60
QY 61 STAMKTENCVAKTLANGTSSMIVPKQKLSASYEKEKELCVKFEQWSESDQVEFEHL 120
DB 61 STAMKTENCVAKTLANGTSSMIVPKQKLSASYEKEKELCVKFEQWSESDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYLKPMLODFITLALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
DB 121 ISOMCHYOHGHINSYLKPMLODFITLALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
QY 181 RYTSQGMMLKILIRMYRTDSLWGLAERGGQYLFRKNRPDGNAPNSFYALYKII 240
DB 181 RYTSQGMMLKILIRMYRTDSLWGLAERGGQYLFRKNRPDGNAPNSFYALYKII 240
QY 241 QDIETIESNMWRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLDNTIKIMDKNTLECK 300
DB 241 QDIETIESNMWRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLDNTIKIMDKNTLECK 300
QY 301 RILGHTGSVLCLOYDERVIITGSSDSTVRYWVDVNTGEMLNTLIHCEAVLHFRNNGM 360
DB 301 RILGHTGSVLCLOYDERVIITGSSDSTVRYWVDVNTGEMLNTLIHCEAVLHFRNNGM 360
QY 361 VTCSKDRSIAVMDMASPTDITLRVLYGHRAAVNVDFDCKYIVASGDRITKVMNTSTC 420
DB 361 VTCSKDRSIAVMDMASPTDITLRVLYGHRAAVNVDFDCKYIVASGDRITKVMNTSTC 420
QY 421 EFVETLNGHKGRIACLOYRDLVYVSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
DB 421 EFVETLNGHKGRIACLOYRDLVYVSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
QY 481 KRIYSGATDGKIKWMDLVLAALDPAPAGTLCRLTVLHSGRVFRLQDFEQIVSSHDDT 540
DB 481 KRIYSGATDGKIKWMDLVLAALDPAPAGTLCRLTVLHSGRVFRLQDFEQIVSSHDDT 540
QY 541 IILWDFLNDPAAQAEPPRSPRTTYISR 569
DB 541 IILWDFLNDPAAQAEPPRSPRTTYISR 569
RESULT 6
Q9P2S8 PRELIMINARY; PRT; 529 AA.
AC Q9P2S8; 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2 ISOFORM B.
GN BTRCP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Katoch M.;
RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,
RT "Molecular cloning and genomic structure of the betaTRCP2 gene on
RT Chromosome 5q35.1.";
RL Biochem. Biophys. Res. Commun. 269:103-109(2000).
DR EMBL: AB033280; BAA92330.1; -
SQ SEQUENCE 529 AA; 60897 MM; 53D181B789313D0E CRC64;
Query Match 79.1%; Score 2400; DB 4; Length 529;
Best Local Similarity 79.3%; Pred. No. 1.5e-191;
Matches 451; Conservative 43; Mismatches 35; Indels 40; Gaps 4;
QY 1 MDPAEAVLQEKALFPMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCTINQETVCLA 60
DB 1 MDPAEAVLQEKALFPMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCTINQETVCLT 60
QY 61 STAMKTENCVAKTLANGTSSMIVPKQKLSASYEKEKELCVKFEQWSESDQVEFEHL 120
DB 61 STAMKTENCVAKTLANGTSSMIVPKQKLSASYEKEKELCVKFEQWSESDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYLKPMLODFITLALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
DB 121 ISOMCHYOHGHINSYLKPMLODFITLALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
QY 181 RYTSQGMMLKILIRMYRTDSLWGLAERGGQYLFRKNRPDGNAPNSFYALYKII 240
DB 181 RYTSQGMMLKILIRMYRTDSLWGLAERGGQYLFRKNRPDGNAPNSFYALYKII 240
QY 241 QDIETIESNMWRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLDNTIKIMDKNTLECK 300
DB 241 QDIETIESNMWRCGRHSIQIRHCRSETSGVYCLQYDDQKIVSGLDNTIKIMDKNTLECK 300
QY 301 RILGHTGSVLCLOYDERVIITGSSDSTVRYWVDVNTGEMLNTLIHCEAVLHFRNNGM 360
DB 301 RILGHTGSVLCLOYDERVIITGSSDSTVRYWVDVNTGEMLNTLIHCEAVLHFRNNGM 360
QY 361 VTCSKDRSIAVMDMASPTDITLRVLYGHRAAVNVDFDCKYIVASGDRITKVMNTSTC 420
DB 361 VTCSKDRSIAVMDMASPTDITLRVLYGHRAAVNVDFDCKYIVASGDRITKVMNTSTC 420
QY 421 EFVETLNGHKGRIACLOYRDLVYVSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
DB 421 EFVETLNGHKGRIACLOYRDLVYVSSGSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
QY 481 KRIYSGATDGKIKWMDLVLAALDPAPAGTLCRLTVLHSGRVFRLQDFEQIVSSHDDT 540
DB 481 KRIYSGATDGKIKWMDLVLAALDPAPAGTLCRLTVLHSGRVFRLQDFEQIVSSHDDT 540
QY 541 IILWDFLNDPAAQAEPPRSPRTTYISR 569
DB 501 IILWDFLNDPAAQAEPPRSPRTTYISR 529
RESULT 7
Q9UKB1 PRELIMINARY; PRT; 542 AA.
AC Q9UKB1; 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE F-BOX PROTEIN FBW1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2 ISOFORM C).
GN FBW1B OR BTRCP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

RN [11]
 RP SEQUENCE FROM N.A.
 RA Genetecell C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M.,
 RA Pagano M.;
 RT "Identification of a family of human F-box proteins.";
 RL Curr. Biol. 9:1177-1179(1999).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FETAL LUNG;
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,
 RA Katoh M.;
 RT "Molecular cloning and genomic structure of the betatrop2 gene on
 RT chromosome 5q35.1.";
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).
 RL EMBL; AF176022; AAF04528.1; -.
 RL EMBL; AB033281; BAA92331.1; -.
 DR INTERPRO: IPR001680; -.
 DR INTERPRO: IPR001810; -.
 DR PFAM; PF00400; WD40; 7.
 DR PFAM; PF00646; F-box; 1.
 DR PRINTS; PR00320; GPROTEINRPT.
 DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
 SO SEQUENCE 542 AA; 62090 MW; 7CD40087EFA5C8A CRC64;

Query Match 78.6%; Score 2384.5; DB 4; Length 542;
 Best Local Similarity 79.0%; Pred. No. 2.9e-190;
 Matches 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;

QY 1 MDPAAVLOEAKLAFNNSSRECCNNGEPPRKIIPEKNSIRQTYNSCARCLNOEYCLA 60
 DB 1 MEP-DSVIEDKTELKCS-----VP-----RSLMLGCANLY--ESMCAL 36
 QY 61 S--TAKTEKNCVAKTKLANGTSSMIVPORKLSASYEKEKELCVKPEOMSEDOYEFVE 118
 DB 37 SCLOSPSVACL---QISNCTSSVIYSKRKPSSEGNQKEKDLCIKTFDQSESDOYEFVE 93
 QY 119 HLISQCHVOHGINSYLRKMLQRFETLAPGLDHAENILSYLDKASICAELVCKE 178
 DB 94 HLISRMCHVOHGINSYLRKMLQRFETLAPGLDHAENILSYLDKASICAELVCKE 153
 QY 179 WYVNTSDGMLMKKLIERMVNTDSLIRGLAERKRGQYLFRKPKPDGNAPNSFYRALYK 238
 DB 154 WQVVISSEGMKMKKLIERMVNTDPLMKGLSERRGWDYLFKNRPTDG--PNSSEFYRSLYPK 211
 QY 239 IIDDIETIESNMRCGRHSIORICRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKTLE 298
 DB 212 IIDDIETIESNMRCGRHSIORICRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKTLE 271
 QY 299 CKRLTGHTGSVYCLQYDEVIITGSSDSYRVWVNTGEMLNTLLHHCBAVLHLRFNNG 358
 DB 272 CLKVLGHTGSVYCLQYDEVIITGSSDSYRVWVNTGEMLNTLLHHCBAVLHLRFNNG 331
 QY 359 MMVTCSDKRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKYIVSASGDRITKYWNTS 418
 DB 332 LMVTCSDKRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKYIVSASGDRITKYWNTS 391
 QY 419 TCEFVRLNKHKRGIACTQYRDRLVYVSGSSDNTIRLMDIECGACRLVLEHBEELVRCIRF 478
 DB 392 TCEFVRLNKHKRGIACTQYRDRLVYVSGSSDNTIRLMDIECGACRLVLEHBEELVRCIRF 451
 QY 479 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCRTLVEHSGVFRLOPFEQIVSSSD 538
 DB 452 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCRTLVEHSGVFRLOPFEQIVSSSD 511
 QY 539 DTLLIMDFLNDPAAQAEPPRSPSRITYTISR 569
 DB 512 DTLLIMDFLNDPAAQAEPPRSPSRITYTISR 542

RESULT 8
 Q9Y4C6 PRELIMINARY; PRT; 550 AA.
 ID Q9Y4C6

AC Q9Y4C6;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)
 DE KIA00696 PROTEIN (FRAGMENT).
 GN KIA00696.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=98403880; PubMed=9734811;
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:169-176(1998).
 RL EMBL; AB014596; BAA31671.1; -.
 DR INTERPRO: IPR001680; -.
 DR INTERPRO: IPR001810; -.
 DR PFAM; PF00400; WD40; 7.
 DR PFAM; PF00646; F-box; 1.
 DR PRINTS; PR00320; GPROTEINRPT.
 DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
 FT NON_TER 1
 SO SEQUENCE 550 AA; 62702 MW; FDE42B5BF3D19C8A CRC64;

Query Match 78.6%; Score 2384.5; DB 4; Length 550;
 Best Local Similarity 79.0%; Pred. No. 3e-190;
 Matches 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;

QY 1 MDPAAVLOEAKLAFNNSSRECCNNGEPPRKIIPEKNSIRQTYNSCARCLNOEYCLA 60
 DB 9 MEP-DSVIEDKTELKCS-----VP-----RSLMLGCANLY--ESMCAL 44
 QY 61 S--TAKTEKNCVAKTKLANGTSSMIVPORKLSASYEKEKELCVKPEOMSEDOYEFVE 118
 DB 45 SCLOSPSVACL---QISNCTSSVIYSKRKPSSEGNQKEKDLCIKTFDQSESDOYEFVE 101
 QY 119 HLISQCHVOHGINSYLRKMLQRFETLAPGLDHAENILSYLDKASICAELVCKE 178
 DB 102 HLISRMCHVOHGINSYLRKMLQRFETLAPGLDHAENILSYLDKASICAELVCKE 161
 QY 179 WYVNTSDGMLMKKLIERMVNTDSLIRGLAERKRGQYLFRKPKPDGNAPNSFYRALYK 238
 DB 162 WQVVISSEGMKMKKLIERMVNTDPLMKGLSERRGWDYLFKNRPTDG--PNSSEFYRSLYPK 219
 QY 239 IIDDIETIESNMRCGRHSIORICRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKTLE 298
 DB 220 IIDDIETIESNMRCGRHSIORICRSETSKGVYCLQYDDQKIVSGLRDNTIKIMDKTLE 279
 QY 299 CKRLTGHTGSVYCLQYDEVIITGSSDSYRVWVNTGEMLNTLLHHCBAVLHLRFNNG 358
 DB 280 CLKVLGHTGSVYCLQYDEVIITGSSDSYRVWVNTGEMLNTLLHHCBAVLHLRFNNG 339
 QY 359 MMVTCSDKRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKYIVSASGDRITKYWNTS 418
 DB 340 LMVTCSDKRSIAVMDASPTDITLRVLYVGHRAAVNVVDEDDKYIVSASGDRITKYWNTS 399
 QY 419 TCEFVRLNKHKRGIACTQYRDRLVYVSGSSDNTIRLMDIECGACRLVLEHBEELVRCIRF 478
 DB 400 TCEFVRLNKHKRGIACTQYRDRLVYVSGSSDNTIRLMDIECGACRLVLEHBEELVRCIRF 459
 QY 479 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCRTLVEHSGVFRLOPFEQIVSSSD 538
 DB 460 DNKRIYSGAIDGKIKYWDVAALDPRAPAGTLCRTLVEHSGVFRLOPFEQIVSSSD 519
 QY 539 DTLLIMDFLNDPAAQAEPPRSPSRITYTISR 569
 DB 512 DTLLIMDFLNDPAAQAEPPRSPSRITYTISR 542

AC 09VME3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, last annotation update)
DE SLMB PROTEIN (SLIMB).
GN SLMB.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zheng Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Flann Koch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Brennan B.P., Bhandari D., Bolshakov S.,
RA Butts K.C., Butman D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pallos B., Delcher A., Deng Z., Duan A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dwyer-Rocha S., Dunkov B.C., Dunn P.,
RA Dubler K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fouts C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatalil M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Kethum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA Renelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY, AND IMAGINAL DISC;
RC Miletich I., Limbourg-Bouchon B.;
RA "Drosophila null slmb clones transiently deregulate Hedgehog-
RT independent transcription of wingless in all limb discs, and induce
RT deapentaplegic transcription linked to imaginal disc regeneration.";
RL Mech. Dev. 93:15-26(2000).
DR EMBL: AE003733; AAF55853.1; -
DR EMBL: AF222924; AAF63314.1; -
DR EMBL: AF222923; AAF63313.1; -
DR FLYBASE: FBgn0023423; slmb.
DR INTERPRO: IPR001680; -
DR INTERPRO: IPR001810; -
DR PFAM: PFO04400; WD40; 7.
DR PFAM: PFO0646; F-box; 1.
DR PRINTS: P000320; GPR0EINBRPT.
DR PROSITE: PS00678; WD_REPEATS; 5.
SO SEQUENCE 510 AA; 58952 MW; F45D5DF126F58A012 CRC64;

Query Match	Best Local Similarity	Matches	Conservative	Score	DB 5	Length	510
67	ENCYVAK-----TKLANGTSMIVPQKOKULASATSEKEKELCVKFFEDMSBDQVFEVPHLIS	122	8	KNVSVKSDLVTSQSPFSSTSSPVCPLNPLS-----HENNRIDILRIDLLA	50		
68.9%	Score 2089.5	DB 5	Length 510				
78.3%	Pred. No. 9.8e-166						
394	Conservative	49	Mismatches	53	Indels	7	Gaps
4	MKTENCVAKTKLANG--TSSMIVPQKOKLAS--YEKEKELCVKFFEDMSBDQVFEV	118					
4	METMDIMETN-SNAQAFETTMLYDPVKRKDDSPYQERELCFOYFTQWSSGOVDVE	62					
119	HLISQMCHYOHCHINSYLAKPMLQGRFITALPARGIDHTIAENLSTLADKSLCAALUYCKE	178					
63	HLISMCHYOHQOINAYLTKPMLQROFTLLPLKIGDHTIAENLSTLADKSLSSSELYCKE	122					
179	WYRVYSDCMWKKLIERVWRTDLSLRGLAERGMQGYLFKNKPPDGNAPNSFYALYPK	238					
123	MLRVISEGMIMKKLIERVYRTDLSLRGLAERNMMQYLFKPRPGQ-TQRPSEFHHELFPK	181					
239	IIIDDIETESNMWGRGHSLSQIRHCSEFTSKGYCCLOYDOKTVSGLRDNTIKWKNLIE	298					
182	IMNDIDSIEHNNMRTGRHMLRINCENSENSKGYCCLOYDQDKTVSGLRDNTIKWRTDQ	241					
299	CKRILTGHTGSLCYQYDERVITGSSOSTYVWVQVNGEMINTLIHCEAVLHLPFNG	358					
242	CVKILMGHTGSLCYQYDOKTVIISSSSDSTVWVQVNGEMINTLIHCEAVLHLPFNG	301					
359	MMVYTSKDRSIAVWMDASPTDITLRVYLVGHRAAVNVYDFDDKYIVSASGDRTIKWMTS	418					
302	MMVYTSKDRSIAVWMDTSPSETTLRRVYLVGHRAAVNVYDFDEKXYIVSASGDRTIKWMTS	361					
419	TCEPFRITNGHKGRIAGCIQYRDRLVYVSSGSDNTIRLMDIEGACACLRVLEGHHELYRCIRF	478					
362	SCEPFRITNGHKGRIAGCIQYRDRLVYVSSGSDNSIRIMDEGCACACLRVLEGHHELYRCIRF	421					
479	DNKRIVSAGADGKIYVMDLVALDGRAPAGTGCTLRTIVESHSGRVERLPQDEFOIYSSSHD	538					
422	DTKRIVSAGADGKIYVMDLVALDGRASNTICLMTIVHTGRVYRLQRPDERFOIYSSSHD	481					
539	DTILIMDFLNDPAAQAEPPRSFS	561					
482	DTILIMDFLNDPENNENKTGRPS	504					
RESULT	12						
09P7V1	PRELIMINARY;	PRT;	506	AA.			
09P7V1	AC						
01-OCT-2000	(TREMBLrel. 15, Created)						
01-OCT-2000	(TREMBLrel. 15, Last sequence update)						
01-OCT-2000	(TREMBLrel. 15, Last annotation update)						
WD-REPEAT	PROTEIN.						
GN	SPAC30.05.						
OS	Schizosaccharomyces pombe (fission yeast).						
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;						
CC	Schizosaccharomycetaceae; Schizosaccharomycetes.						
OX	NCBI_TaxID=4896;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RA	STRAIN-972H-;						
RA	McDougal R.C., Rajandream M.A., Barrell B.G., Brown S., Murphy L.,						
RL	Jones L., McNeil A., Harris D.;						
RL	Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.						
DR	EMBL: AL136538; CAB66464.1; -						
SEQUENCE	506 AA; 58256 MW; CER34DAEFFBCE10 CRC64;						
Query Match	22.7%;	Score 690;	DB 3;	Length 506;			
Best Local Similarity	30.4%;	Pred. No. 3.1e-49;					
Matches	163;	Conservative	104;	Mismatches	167;	Indels	102;
							Gaps
67	ENCYVAK-----TKLANGTSMIVPQKOKULASATSEKEKELCVKFFEDMSBDQVFEVPHLIS	122					
8	KNVSVKSDLVTSQSPFSSTSSPVCPLNPLS-----HENNRIDILRIDLLA	50					


```

OY 123 OMCHYOHGHINSYLPMDORPEITIALPARGDHTAENILSYLDAKSISCAEYLCKEYRY 182
Db 51 SLSGEGVAAYVNNHRSLLFTDFTEVP---EEVSLVFSTYLDLDCCKCKLMSKRKRL 106
OY 183 TSDGMLMKLL-----ERMYTDSLWRG-----LAERGMG----- 213
Db 107 LEDPGIMATLYOKGMFVENVULNEFEAMRYTHKFPQPRENFLLQONIGPYGTMLPQ 166
OY 214 OYLFRKNRPDQNAAPNSFYRLALYKTIQDIETIESMNRGRHSLOIRHCS----- 264
Db 167 QPFI-----DSNGRPLMWSTLY---KEAHNLDMSNRHGRPLVSTFNPSIRFPADQF 217
OY 265 -ETSGKVYCLQYDQCKIYSGLDYTIKIKNPLECKRIILTGHTSGVYLCQYDER--VII 321
Db 218 RATIDSVYCVQYDDEIMVSGSKDPTSVWMDVNSFLLIKYLHGSGVSLCDFCRRRLIV 277
OY 322 TGGSDSYRVWDVNTGEMKNTLIHHCZAVLHLFRNNGMAYTCSKDRSIAYW--DMASPTD 379
Db 278 SGSSDSTIIMDMONRRPLKYFGHTNVLGVVVSSENYIISSSDHTRAPRWRLDAPSPAE 337
OY 380 ITLRLVGLHRAAVVYVDFQCK--YIVASADRTIKYWNSTQCEPFTVLNGHKGRTACLO 437
Db 338 ACM-HVLMGHLASVNSQVYSKTLIYTASSDRLLKRWMDITTGACITIIHAHGRIGLCAO 366
OY 438 YRDLRVYSGSSDNTIRLMDIEGCACLRVLEEGHEBLVRCIFRDNKRIYSGAYDKIKYMDL 497
Db 397 YNGKFIYVSGSSDLTRIPEASGKILMLQGHEDLRTIYFBNDEK IYSGGYDGTVRIMN- 455
OY 498 VAALDRAPRATLCLRTLYEH-----SGVRFLODDEQOYVSSSDHDTIILMDF 546
Db 456 -----FNGEQHCYVLIHNSRNSYVPELODPRHRRIIACHTSHSELVWNE 497

```

ID	Q9VZF4	PRELIMINARY;	PRT;	1326 AA.
Q9VZF4	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	CG15010 PROTEIN.			
GN	CG15010.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wotman J.R., Yancelli M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazesi R.G., Champe M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abtill J.F., Agbayani A., An H.-J., Andrews-Plankkoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,			
RA	De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbatt W.M., Glasser K.,			
RA	Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.H., Houck C.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbewam C.,			
RA	Jaitai M., Kalush F., Kaspen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			

RA Liu X., Attwater B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Malshina N.V., Mobbart C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,"
RT "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195(2000).
EMBL: AE003480; AAF47869.1; "-
FLYBASE: FBgn0035516; CG15010.
DR INTERPRO: IPR001680; "-
DR INTERPRO: IPR001810; "-
DR PFAM: PF004400; WD40; 7.
DR PFAM: PF00646; F-box; 1.
DR PRINTS: PR00320; GPOTFEINBRPT.
DR PROSITE: PS00678; WD_REPEATS_5.
SQ SEQUENCE 1326 AA; 141360 MW; 3F42C873CFA3027F CRC64;

Query Match	21.8%;	Score 661;	DB 5;	Length 1326;
Best Local Similarity	34.9%;	Pred. No. 3.1e-46;		
Matches 158;	Conservative 76;	Mismatches 167;	Indels 52;	Gaps 12

OY	105	FEQSDSEDOVEFVHLTHISOMCHONHGINSLYLPMLQDREFTLTPARGLDIAENLSTYL	164
Db	852	FQKSHVERLLADRLDLHDHDPQOVRRIMMKVIERPQDROFISLPR---ELALVYLSL	907
OY	165	DAKSLCAAELYSCKEWYRVVYSDGMLMKLLERMYRTDLSLRGAEERRGMQGYLFKNKP--	222
Db	908	EPKDLRAAOTCRSMRFLCDNLMLWEKCKRA-----QILAEPR-----SDRPKRG	953
OY	223	-DGNAPP--NSFYRLAYKIKIIDIETTESNRRCGRHSIORLHCSEFSKSG----UYCLOY	275
Db	954	RDGNMPRLPASAQWKAAYNR--QHI--IEMNR-----SRPVKRPYLLKQHDHIVTICLOF	1002
OY	276	DDQKIVSGLRDNTIKIMDKNLTLECKRLIHGHTGSVYCLQYDERAVITITGSSDSTVKRYMDN	335
Db	1004	SGNRIVSGSDNTLTKVNSAVANGKCLRLVHGHTGGVWSQMSGNIITISGTDRTFLKWMDD	1065
OY	336	TGERMLNTLIHHCVALYHLIRFENNOMAYTCSKDRSIAMDNASPTDILRLVYLGHRAVAVY	395
Db	1064	SGACVHTLQHTSTVYRCMHHLGSKVYSGSDATLRFWMDIEQSCSL---HYLVGHILAAVNC	1120
OY	396	VDEPDKITVYASGDRITKAVNNTSTCEPFRYTLNGHKAGICLQDRDLVYSGSSDNTIRIM	455
Db	1121	VQYUGKLTIVSAGYAWYKIMHPEROECLHNTLQHTTRVYSLDFQDLHVVYSGSLDLSIRFW	1180
OY	456	DIEGCACRLVLEGHEELVRCIRDRNKRIYSGAIVGADGKIKYWDVLVAALDRPAPAGTICLRTL	515
Db	1181	DVEITGNCKHTLMHGQSLTSGMELRQNLIVSGANDSTVKWMDITTG-----QCLQTL	1233
OY	516	V--EHSGRVFRLODFEQIYVSSSHDDTILIMD	545
Db	1232	SGPRKHHSAVTCLQFNSRFVVTSSDSDGTAKLMD	1264
RESULT	14		
Q9NDX6			
Q9NDX6	PRELIMINARY;	PRT;	553 AA.
AC	Q9NDX6;		
DT	01-OCN-2000 (Tremblrel, 15, Created)		
DT	01-OCN-2000 (Tremblrel, 15, last sequence update)		
DT	01-OCN-2000 (Tremblrel, 15, last annotation update)		
DE	CDNA FLJ11071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10 PROTEIN.		

THIS PAGE BLANK (USPTO)